ROLE OF APOBEC3 ENZYMES IN HIV-1 EVOLUTION

A Thesis Submitted to the

College of Graduate and Postdoctoral Studies

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in the

Department of Biochemistry, Microbiology, and Immunology

University of Saskatchewan

By

Nazanin Mohammadzadeh

PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirement for a Postgraduate degree from the University of Saskatchewan, I agree that the libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done.

DISCLAIMER

References in this thesis to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions expressed herein do not state or reflect those of the University of Saskatchewan, and shall not be used for advertising or product endorsement purposes. Requests for permission to copy or to make other use of the material in this thesis, in whole or in part should be addressed to:

Head of the Department of Biochemistry, Microbiology and Immunology
University of Saskatchewan
107 Wiggins Road
Saskatoon, Saskatchewan S7N 5E5
Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada

ABSTRACT

There are approximately 36.7 million people living with HIV while 1.8 million become newly infected every year. HIV belongs to the Retroviridae family and is an enveloped, positive sense (+) single stranded RNA virus. HIV has two subtypes, HIV-1 and HIV-2. HIV-1 is responsible for 95% of infections worldwide and is the focus of this thesis. There has not been any report on a conventional cure of this disease and the diagnosed HIV+ individuals need to receive lifelong treatment with antiretroviral drugs. Besides the antiretroviral drugs, there are host restriction factors that fight the infection in addition to the conventional immune barriers and responses.

The APOBEC3 (A3) family of enzymes are part of an intrinsic immune system in humans and can act as host restriction factors to restrict the replication of HIV in CD4+ T cells by deaminating cytosine to uracil on the (-) DNA of HIV during reverse transcription. This promutagenic activity, if it occurs frequently enough, can cause viral hypermutation and inactivation. The A3 family contains seven members (A3A to A3H, excluding A3E) some of which developed gene variations that result in protein polymorphisms due to selective pressure over evolutionary time. These polymorphisms acquired enhanced antiretroviral activity to fight off HIV. To combat this, HIV has an accessory protein termed virion infectivity factor (Vif) that interacts with A3 enzymes and components of a Cullin5 E3 ligase complex. This complex causes ubiquitination and degradation of A3 enzymes in HIV infected host cells.

HIV also has a high level of genetic variation that allows the virus to escape immunological and pharmacological barriers. Besides the lack of proofreading activity of reverse transcriptase, viral recombination and high rate of replication, the A3 enzymes have a potential role in increasing the genetic diversity of the HIV replicating pool if Vif mediated degradation is not complete. As a result, human A3 enzymes can contribute to this diversity directly by causing sublethal mutations and indirectly when the lethally mutated RNA genomes are “rescued” after virus recombination. While A3 enzymes may have this potential, and some research has found that A3 enzymes cause drug resistance mutations, other research groups have reported that in the presence of Vif the mutation rate of A3s is lower than the mutation rate of reverse transcriptase and their contribution to HIV evolution is not significant.

To identify all the contributors to genetic variation of HIV and have a reliable answer for the current controversy in the field we aimed to determine the mutation rate of A3s in wild-type HIV-

1 infection and investigate if A3-induced mutations can yield drug resistant HIV-1 variants. Moreover, we aimed to study the restriction abilities of A3F allele variants carrying a single nucleotide polymorphism (SNP). We chose the SNP that causes the A3F 231I/V polymorphism because most people carry the heterozygous alleles.

To test if A3 enzymes could contribute to HIV-1 evolution, we established a system to isolate HIV-1 drug resistant variants using peripheral blood mononuclear cells (PBMCs). We then used this method for a continuous spreading infection on infected U87 CD4+ CXCR4+ cells with a HIV-1 that was or was not exposed to A3G and A3F in a single replication cycle to discern the contributions of A3-mediated mutagenesis. The viruses were grown in the presence of increasing amounts of antiretroviral drugs. Viruses were collected and tested for viability using a reverse transcriptase activity assay. The data from four independent experiments with different drugs showed that A3G and A3F can induce drug resistance more times than when the virus was grown in their absence. However, drug resistant viruses exposed to A3G and A3F had less replication capacity than drug resistant variants that arose from HIV reverse transcriptase errors alone. Integrated proviral DNA and HIV genomic RNA (converted to cDNA) were sequenced to analyze A3-induced mutations. Importantly, for the first time we have quantified the amount of A3-induced mutations that occurs in the presence of Vif and shown that A3-induced mutations are approximately 4-fold above the background mutation level of HIV-1 reverse transcriptase alone. We also conducted similar experiments in CEM CD4+ T cells where endogenous A3 enzymes are expressed and obtained similar results.

It is known that A3s have numerous polymorphisms which can affect antiretroviral activity of these enzymes. Namely, the CEM cells used in our experiments and numerous other labs have heterozygous alleles for A3F. As a result, we wanted to examine if the resulting A3F polymorphisms can be an additional factor that effects A3 activity in the presence or absence of Vif. We found that the A3F 231V polymorphism, in comparison to the A3F 231I, resulted in higher expression and virus encapsidation in the presence and absence of Vif. Specifically, we observed approximately 4-fold more virus restriction and mutations in proviral DNA.

Altogether, these data suggest that A3 enzymes can be a viable option as a therapy for HIV-1 if the interaction between Vif and A3s were disrupted. However, such a potential therapy would also depend on if the A3 genotype of the person contained highly active variants, such as A3F 231V.

TABLE OF CONTENTS

ACKOWLEDGEMENTS

I would like to thank my supervisor, Dr. Linda Chelico, for the mentorship, patient guidance, caring support and kind understanding she has provided throughout my time as her student. I have been extremely lucky to have a supervisor who is both an excellent scientist and a great human being. Linda's strong character as a successful woman will be a light on my path in science and in life. Her passion for science and research has inspired and influenced me. I want to thank Linda for an amazing environment that she provided for us to learn, be curious, and discover.

Secondly, I would like to express overwhelming thanks to Dr. Mirek Cygler and Dr. Wei Xiao for their input and Dr. Harold Bull for his insightful questions. Their valuable advice and constant support throughout my study that has allowed my research and scientific knowledge to grow. I would also like to thank the Department of Biochemistry, Microbiology & Immunology and for their support.

To the members of the lab past and present, that I shared most of my time in Canada with, my office mates Lai, Mo and Tyson that made me talk and smile when I didn't feel like it. Also, Madison and Anjuman that cared about my research and answered my questions so promptly.

An important part of my acknowledgments belongs to Robin, my best friend and partner. He was beside me and there for me in such a heartwarming way that was beyond my expectations. His work ethic, moral and generosity of spirit has taught me a lot and I'm honored to have worked with him.

DEDICATION

To my mom, زهره

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

(-) DNA	Minus strand/template strand
(+) DNA	Positive strand/coding strand
3'UTR	3' untranslated region
3TC	Lamivudine
A3	APOBEC3
A3H hap I	APOBEC3H haplotype I
AID	Activation induced cytosine deaminase
AIDS	Acquired immunodeficiency virus
AIDS	Acquired immunodeficiency syndrome
APE	Apurinic/apryidiminic endonuclease
APOBEC	Apolipoprotein B mRNA editing, catalytic-enzyme
APOBEC3	Apolipoprotein B mRNA editing, catalytic-enzyme polypeptide like 3
ART	Anti-Retroviral therapy
ARV	Anti-retroviral drugs
ATV	Atazanavir
AZT	Azidothymidine
BER	Base excision repair
BNAbs	Broadly neutralizing antibodies
BST-2	Bone marrow stromal antigen 2
CA	Capsid
CBFβ	Core binding factor
cDNA	Complementary DNA
co-IP	Coimmunopercipitation
cPPT	Central poly purine tract
CRF	Circulating recombinant form
CTD	C-terminal domain
CTL	Cytotoxic T lymphocytes
Cul5	Cullin 5
dNTP	Deoxyribonucleoside triphosphate
EFV	Efavirenz
EloB	Elongin B
EloC	Elongin C
Env	Envelope
ESCRT	Endosomal sorting complexes required for transport
FIs	Fusion inhibitors
FTC	Emtricitabine
Gag	Group specific antigen
GP	Glycoprotein
gRNA	Genomic RNA
GUD	Genital ulcer disease
HIV	Human immunodeficiency virus
HIVDR	Human immunodeficiency virus drug resistance
IN	Integrase

LPV/r	Lopinavir/ritonavir
LTR	Long terminal repeat
MA	Matrix
MHC	Major histocompatibility complex
mRNA	Messenger RNA
NC	Nucleocapsid
Nef	Negative factor
NLD	Nuclear localization signal
NNRTIs	Non-nucleoside reverse transcriptase inhibitors
NRTIs	Nucleoside reverse transcriptase inhibitors
NTD	N-terminal domain
OIs	Opportunistic infections
ORF	Open Reading Frame
PBMC	Peripheral blood mononuclear cell
PBS	Primer binding site
PCR	Polymerase chain reaction
PEP	Post-exposure prophylaxis
PIC	Pre-integration complex
PIs	Protease inhibitors
Pol	Polymerase
PPT	Poly-purine tract
PR	Protease
PrEP	Pre-exposure prophylaxis
Prot	Protease
RBX2	Ring box protein 2
RNase A	Ribonuclease A domain enzyme
RNase H	Ribonuclease H domain enzyme
RT	Reverse transcriptase
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIV	Simian immunodeficiency virus
SIVcpz	Simian immunodeficiency virus from chimpanzees
SNP	Single nucleotide polymorphism
SOCS2	Suppressor of cytokine signaling 2
ssDNA	Single-stranded DNA
SU	Surface
TAR	Trans-Activation Response element
Tat	Transcriptional transactivator
TDF	Tenofovir disoproxil fumarate
TLE	Tenofovir disoproxil fumarate, Lamivudine and Efavirenz
TM	Transmembrane
UDG/UNG	Uracil DNA glycosylase
URF	Unique recombinant form
VCBC	Vif/ CBFβ /EloB/C complex
Vif	Viral infectivity factor
VLP	Viral like particles
Vpr	Viral protein r

Vpu	Viral Protein u
Vpx	Viral protein x
VSV-G	Vesicular stomatitis virus glycoprotein
WHO	Worlds health organization
WT	Wild type
Z1	Zinc binding domain type 1
Z2	Zinc binding domain type 2
Z3	Zinc binding domain type 3
ZDV	Zidovudine

1.0. INTRODUCTION

1.1 HIV

Human Immunodeficiency Virus (HIV) is grouped into the genus Lentivirus and the family of Retroviridae. It has emerged from Africa because of zoonotic transmissions of simian immunodeficiency virus (SIV) from primates to humans in the early 1900s due to injuries during meat butchering practices of hunted monkeys and chimpanzees[1]. Epidemic origins of the infection were tracked down to the 20th century in the Congo/Central Africa as a result of colonialism leading to cultural changes, such as, increased sex work and genital ulcer disease (GUD) outbreaks[2]. Independent transmission of SIV from chimpanzees and gorillas originated HIV-1 and its subgroups, M (main), O (outlier), N (non-M or non-O) and P (putative). Group M that is responsible for 95% of infections worldwide is divided into nine phylogenetic subtypes or clades (A, B, C, D, F, G, H, J, and K)[3]. Subtypes A, B, and C are the most prevalent forms of HIV-1 worldwide where subtype C accounts for 50% of all HIV-1 infections[4]. Cocirculation of different subtypes in a population increases the chance of inter-subtype recombination. This recombinant virus is called a unique recombinant form (URF). When URFs are identified in at least three epidemiologically unlinked people, they are classified as circulating recombinant forms (CRFs). URFs and CRFs increase the chance of superinfection resulting in increased genetic variation of the replicating population of the virus[3,4]. The less pathogenic and transmissible type of HIV, HIV-2, is limited to small regions in South Africa, Europe, India, and the United States[5]. HIV-2 contains 8 groups (A–H), and is descended from sooty mangabeys[6]. Lower transmission rate of HIV-2 is correlated with lower viral load. Usually the HIV-2 viral load is approximately 37-fold lower than the viral load in HIV-1 infected individuals and appears to be responsible for the lower mother-to-child transmission levels, rather than factors inherent to the virus[7].

1.2 AIDS

HIV is a major global health problem, which has taken over 35 million lives so far. According to the WHO approximately 36.7 million people are living with HIV as of 2016. HIV is transmitted from exposure to body fluids of infected individuals (i.e. blood, breast milk, semen and vaginal secretions). The virus targets the immune system and causes depletion of CD4+ T cells and loss of their function which is the underlying cause of gradual immunodeficiency in infected

individuals (normal range for CD4+ T cell count is 500-1,500 cells/ml and <350 is immune deficient)[8]. The most advanced stage of HIV infection is Acquired Immunodeficiency Syndrome (AIDS), which can take from 2 to 15 years to develop depending on the individual. This syndrome is characterized by the development of rare opportunistic infections (OIs) and/or cancers that people with healthy immune systems normally suppress[9]. Besides the pathological outcomes, HIV/AIDS has a great socioeconomic impact on the countries and individuals. Persons infected with the virus must not only cope with the disease but also, they may suffer from discrimination. Stigmatization against HIV/AIDS is associated with fear of death (instrumental stigma), symbolic association with homosexuality and injection drug use (symbolic stigma) and discrimination against people that are connected to HIV/AIDS (e.g., volunteers, children of parents living with HIV)[10,11].

1.3 Genomic Structure and Function

The HIV virion contains two copies of single stranded positive-sense RNA molecules with a 5′ cap and 3′ poly(A) tail. The 9719 bp genome of the virus contains nine genes that encode fifteen viral proteins through three ORFs that are further categorized into three main groups: structural proteins, regulatory proteins and accessory proteins[9,15]. The 5′ long terminal repeat (LTR) region codes for the promotor and is followed by the *gag (group-specific antigen)* gene. The *gag* precursor is also known as assemblin because it plays a role in viral assembly. Gag is a 55 kDa myristoylated protein, referred to as p55, that digests into matrix proteins (MA, p17), the capsid protein (CA, p24), the nucleocapsid (NC, p7) and a small Pro-rich protein (p6) by the viral protease[12,13]. Ribosome frameshifting near the 3' end of *gag* produces the *gag-pol* precursor. This precursor is subjected to cleavage by protease into reverse transcriptase/RNase H (RT, p66 and p51), protease (PR, p12), RNase H (p15) and integrase (IN, p32)[9,14]. In addition to *gag* and *pol*, HIV contains another structural protein, env. The env precursor (glycoprotein (gp)160) contains the gp120 (surface protein, SU) and gp41 (transmembrane protein, TM) glycoproteins that will reside on the external surface of the viral particle[14] (Figure 1.1). HIV-1 and HIV-2 are similar in terms of structural and regulatory proteins. However, they have different accessory proteins[6,9,15,16]. Accessory proteins are used to antagonize host restriction factors. Host restriction factors are defined as host proteins that can specifically block a replication step of the virus[17]. Both HIV types contain the regulatory proteins, Tat and Rev.

1.3.1 Transactivator protein and RNA splicing-regulator

Tat (transactivator protein) activates transcription of viral genes and regulates latency of the virus by binding to the host transactivation response (TAR) element and activating transcription initiation and elongation from the LTR promoter[18]. The RNA splicing-regulator, Rev, is a functionally conserved phosphoprotein. Rev cycles between the nucleus and cytoplasm but is primarily localized in the nucleus. This protein regulates the nuclear export of non-spliced and partially spliced viral mRNAs[19,20]. Besides the structural and regulatory genes, HIV has a set of accessory proteins that increase the pathogenesis of the virus and improve the viral fitness by counteracting the host antiviral responses. The accessory proteins Vif, Nef, Vpu, Vpr and Vpx lack any enzymatic activity. They interact with cellular factors to mediate degradation of host restriction factors, change their localization or redirect their normal function. It has been suggested that the function of accessory proteins has evolved as a result of the selective pressures of continual replication in primate hosts and interacting with host restriction factors[21,22,9].

1.3.2 Viral infectivity factor

Vif (Viral infectivity factor) is an essential factor for the production of infectious virions *in vivo* and is found in all Lentiviruses, except equine infectious anemia virus. Studies have shown that Vif-deficient (ΔVif) viruses can replicate unaffected in many tissue culture systems. However, there is no report on clinical replication competent Δ*vif* HIV isolates[23]. HIV-1 Vif is a cytoplasmic protein and targets host restriction factor APOBEC3 (A3) enzymes. A3s are deoxycytidine deaminases of single-stranded (ss) DNA that catalyze C to T mutations on the (-) DNA of the proviral genome during reverse transcription. Since my thesis investigates the interaction between HIV and A3 enzymes I will elaborate on HIV-1 Vif in more detail than other HIV proteins.

The Vif protein mediates the polyubiquitination and subsequent degradation by the 26S proteasome of the five HIV restrictive A3 proteins (A3C S188I, A3D, A3G, A3H and A3F)[23-26]. Vif recruits an E3 ligase complex consisting of Cullin 5 (Cul5), Elongin B (EloB), Elongin C (EloC) and Rbx2 to mediate A3s degradation. Primate lentiviral Vif proteins also require Core Binding Factor b (CBF-β) for stability. Although Vif can only associate with the E3 ligase in the presence of CBF-β, the CBF-β is not a functional part of the E3 ligase complex. Importantly, CBF-

β permits the biochemical purification of the tetrameric complex (CBF-β, Vif, EloC and EloB)[27-32]. Vif by means of two domains, organizes the formation of Vif, CBF-β, EloC, EloB and Cul5 pentameric complex (Figure 1.2). The larger domain (α/β domain) of Vif (red) binds to CBF-β (blue) (Figure 1.2). The α/β domain contains a five stranded antiparallel β sheet with three α-helices that are tightly packed. The N-terminal region of Vif makes hydrophobic contact with residues on CBF-β. Vif is charged positively on its surface and this binding orientation keeps the positively charged surface exposed. The smaller domain contains two α-helices (α3 and α4) with a Zinc finger motif (H_{108}-C_{114}-C_{133}-H_{139}) between these two domains. The α3 binds to Cul5 (orange) and shares some residues with Zinc finger motif in this binding (Figure 1.2). The α4 harbors the BC-box motif that is required for interaction with EloC (Figure 1.2, yellow). Connection of EloB (magenta) to this hetero-oligomer is only through binding to EloC (Figure 1.2)[33]. Vif also directly binds to the selected A3s and induces their ubiquitination which results in exclusion of the A3s from the cytoplasm. This decreased quantity effects the encapsidation efficiency of A3s in HIV-1 virions, which is an important step in restrictive activity of the A3s.

1.3.3 Negative regulatory factor

Nef (Negative regulatory factor) is expressed only in primate Lentiviruses and at the early stage of viral replication. Studies have shown that patients carrying Nef deficient viruses have delayed onset of AIDS as a result of long-lasting low level of viral replication[34-36]. Nef is a myristoylated protein that associates with plasma membrane and prenuclear membrane complexes. Nef is a multifunction protein that acquires its main activities as a result of connection to cellular vesicular trafficking machinery to disturb cell signaling[37,38]. Nef induces down-regulation of cell surface receptor CD4 from the plasma membrane by increasing its uptake into the endosome-lysosome compartment. This protects the infected cells from cytotoxic super infection[39]. Nef hides the infected cells from the host immune cells by down regulating major histocompatibility complex class I (MHC-I) (reviewed in ref.[40]). Nef also causes T cell receptor (TCR) pathway dysfunction and promotes downregulation of CD28 receptor which is believed to be involved in the latency of the virus[41]. Nef directly binds to SERINC5 (Serine incorporator 5), a host restriction factor, and down regulates its cell surface expression via endosome-lysosome pathway. SERINC5 increases susceptibility of the Env to broadly neutralizing antibodies (BNAbs)[42,43].

1.3.4 Virus protein unique

Vpu (Virus protein unique) is only encoded in HIV-1. Vpu is a multifunction protein. It down regulates cell surface receptor CD4 in a multi-step process via ubiquitin-proteasome pathway[44]. Vpu facilitates release of viral particles through a Vpu ion channel activity[45,46]. Moreover, down regulation of CD4 increases the release of fully infectious virions. Vpu antagonizes the host restriction factor tetherin or BST-2 (Bone marrow stromal antigen 2)[47]. Upregulation of tetherin/BST-2 is triggered by interferon response caused by viral infection. Tetherin/BST-2 targets conserved regions of HIV-1 and inhibits the release of enveloped virus particles from infected cells[47,48]. Vpu directly interacts with tetherin and degrades it in a ubiquitin-dependent endosomal pathway[49]. HIV-2 does not encode for Vpu but is able to antagonize tetherin using Env glycoproteins. SIV utilizes Nef protein to combat tetherin induced challenges[50,51].

1.3.5 Virus protein r

Vpr (virus protein r) is encoded in HIV-1 and HIV-2. Vpr is localized to the nucleus and is involved in transport and nuclear localization of the pre-integration complex (PIC). This assures replication of the virus in non-dividing cells like monocytes and macrophages. Vpu induces G2/M cell cycle arrest in proliferating cells[52,53,54]. G2/M cell cycle arrest and subsequent apoptosis enhances pathogenicity of the virus by providing additional time for virion packaging and protein synthesis[55,56]. Vpr blocks the cell cycle at G2 to mitosis transition and escapes from innate immune sensing by recruiting the host cellular factor SLX4 complex (SLXcom)[55]. SLX4com is an endonuclease involved in the repair of double stranded DNA breaks and stalled/collapsed replication forks by homologous recombination[57,58]. Vpr recruited SLXcom degrades excess HIV-derived nucleic acids which can trigger the innate immune response upon accumulation (reviewed in ref[59]).

1.3.6 Virus protein x

Vpx (Virus protein x) is unique to HIV-2 and SIV. Vpx is very similar to Vpr structurally and functionally and it is a product of gene duplication events in Vpr[60]. Vpx antagonizes the host restriction factor SAMHD1[61,62]. SAMHD1 restrict replication of HIV during reverse transcription step by decreasing the levels of cellular dNTP pools[63,64].

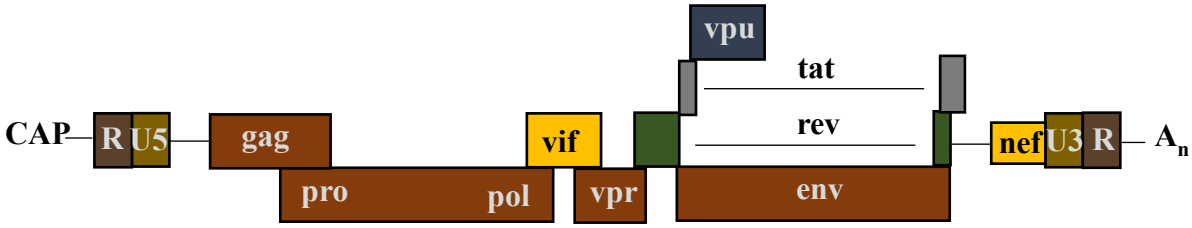


Figure 1.1 Schematic model of the HIV-1 full-length RNA genome. HIV-1 linear double-stranded proviral DNA contains structural genes gag, pol, and env coding for matrix (MA), capsid (CA), Pro-rich protein, nucleocapsid core proteins (NC), protease (PR), reverse transcriptase (RT), RNase H, surface subunit glycoprotein (SU), and a smaller transmembrane protein (TM) and integrase (IN). HIV has six accessory gene products: Tat, Rev, Vif, Nef, Vpr, and either Vpu in HIV-1 or Vpx in HIV-2. The mRNA produced contains a CAP and poly A tail. Adapted from[16].

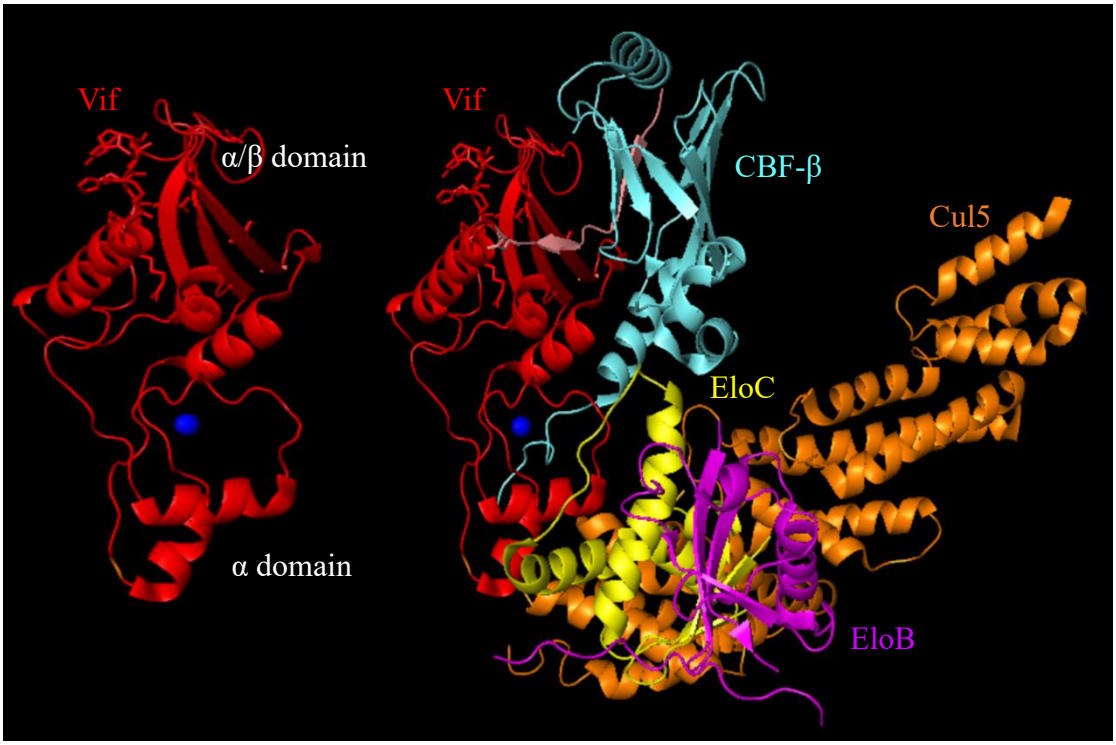


Figure 1.2 Structure Vif-CBF-B-CUL5-ELOB-ELOC complex. U-shaped structure of the Vif-organized Vif–CBF-β–CUL5–ELOB–ELOC complex. Overall structures of Vif–CBF-β–nCUL5–ELOB–ELOC in two different orientations. Colour codes for the proteins are indicated. The dark blue spheres indicate Zn. Adapted from[33].

1.4 HIV Replication Cycle

The infection begins with the binding of gp120 to the CD4 receptor on T helper cells, macrophages, dendritic cells and astrocytes. This binding results in conformational changes and stimulates binding of gp120 to chemokine receptors (CCR5, CXCR4 or both). Binding of gp120 to CD4 and to the co-receptor triggers an additional conformational change in gp41(TM). Attachment of viral particle to a host cell takes around 30 minutes to 2 hours. The gp41 protein subsequently forms a channel into the plasma membrane of the target cell resulting in fusion of cell membrane to viral envelope and viral entry (Figure 1.3, steps 1 and 2)[6,15,65].

Entry of the capsid into the cytoplasm activates the reverse transcription process. The host deoxynucleotide triphosphates (dNTPs) diffuse into capsid enabling DNA synthesis[66,67] . During this process HIV (+) viral RNA is converted into a double stranded DNA (Figure 1.3, step 4). HIV packages two RNA genomes, resulting in a diploid genome, although only one proviral DNA is synthesized. The diploid genome benefits the virus by recombination through template switching at the DNA level and increasing the genetic diversity[68] . The viral RNA is flanked at both ends by regulatory sequences (R-U5 at 5′ and R-U3 at 3′). During the reverse transcription the RNA dependent DNA polymerase part of reverse transcriptase uses a host $tRNA_{Lys3}$ to synthetize a complementary strand of RNA (Figure 1.4). This newly synthesized (–) DNA will transfer from the 5′ end to the 3′ end through NC nucleic acid chaperone activity resulting in identical copies of LTRs containing the U3-R-U5 sequences at the ends of (–) DNA (Figure 1.4). The RNA strand of the resulting RNA-DNA hybrid is the substrate for RNase H. The RNase H activity of RT results in endonucleolytic cleavage of the genomic RNA, leaving the newly synthesized (–) DNA exposed. The HIV RT is a heterodimer containing the p66 portion with polymerase and RNase H functions and the p51 portion that is required for stability[66,67] . However, there are two purine-rich patches in the RNA, called the polypurine tracts (PPT) that are resistant to the action of RNase H for longer[68]. As a result, the PPTs serve as primers for synthesis of the (+) DNA. Once the 18 nucleotides of tRNA have been copied into the 3′ end of the (+) DNA, RNase H cleaves the tRNA and these exposed 18 nucleotides that are complementary to the Primer Binding Sequence (PBS) on the (–) DNA will anneal, resulting in the second strand transfer. Extension of the plus and minus strands by RT leads to the synthesis of the complete double-stranded linear viral DNA[2,4,6,9,66,69]. In order to integrate into the host genome HIV forms a pre-integration complex (PIC) which is

responsible for nuclear import. The PIC consists of the viral DNA, IN, RT, MA, and Vpr. MA and Vpr both contain nuclear localization signals (NLS) that are important for nuclear import (reviewed in ref[70]). After uncoating and removal of the capsid, the PIC, is transported to the host nucleus via nucleopores (Figure 1.3, step 5). It is believed that viral and cellular factors modulate the stability of the capsid. However, the exact location and time frame of uncoating process is yet to be determined (Figure 1.3, step 3)[71].

After integration (Figure 1.3, step 6), the viral DNA (now called proviral DNA) is replicated along with the host genome. Transcription (Figure 1.3, step 7) initiates from the U3 promoter upstream LTR and requires the Tat protein for efficient elongation. Tat recruits cellular factors to TAR element resulting in activating transcription initiation by DNA-dependent RNA polymerase (RNA Pol II) and elongation from the LTR promoter (reviewed in ref[65]). Transcription results in production of completely spliced small mRNAs which can independently get exported from the nucleus, or unspliced and singly spliced mRNAs that require Rev for nuclear export (Figure 1.3, step 8) (reviewed in ref [65]). The completely spliced mRNAs serve as templates for Tat, Rev and Nef. Incompletely spliced mRNAs encode Env, Vpu, Vif and Vpr. Unspliced viral RNAs will incorporate into viral particles that will bud off the cell after assembly (Figure 1.3, step 9)[72].

HIV gag polyprotein (and its mature cleaved structural proteins, CA, MA and NC) is required for virus like particle (VLP) assembly and budding (Figure 1.3, steps 10 and 11). Gag associates with the plasma membrane and initiates all of the essential events for viral assembly like packaging the genomic RNA, creating spherical particles by protein-protein interaction and concentrating the Env protein. These assembling viral packages contain two copies of the viral RNA, cellular $tRNA^{Lys,3}$, Env, Gag, PR, RT and IN[73,74]. HIV uses host trafficking machinery ESCRT (endosomal sorting complexes required for transport) for budding and virion release (Figure 1.3, steps 11 and 12) (reviewed in ref [73,75]). Released virions undergo protease-mediated maturation during or after budding. Protease catalyzes lysis of the precursor peptides Gag and Gag–Pol to produce MA, CA and NC, and the enzymes PR, RT, and IN. Mature virions are infectious and will target other CD4+ cells (Figure 1.3, step 13)[6,15,65,76].

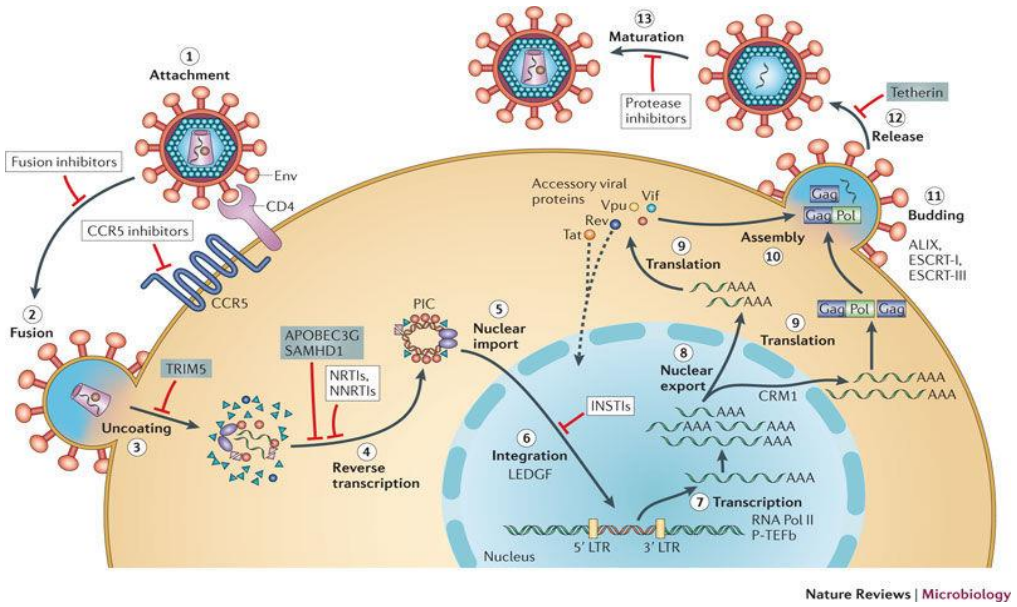


Figure 1.3 Overview of HIV replication cycle. Step 1: Attachment of the envelope glycoprotein spikes (gp120 and gp41) to CD4 receptor and co-receptors this strep leads to **step 2**, fusion of virus and the cell membrane and entry of the virus particle into the cell. **Step 3**, uncoating. **Step 4**, reverse transcription and production of pre-integration complex (PIC). In the next step PIC gets imported into the nucleus **(step 5),** PIC-associated IN directs formation of the integrated provirus **(step 6).** After proviral transcription **(step 7)** viral mRNAs and unspliced RNAs are produced. The larger amount of which require energy-dependent export to leave the nucleus **(step 8).** Some mRNAs are used as a template for protein production **(step 9)** then full length RNAs and other expressed proteins are assembled into budding virions (**step 10**). The virion will bud off the cell through mediation of ESCRT **(step 11**) and release **(step 12)** from the cell. Then the virion will undergo PR-mediated maturation in **step 13** to create an infectious viral particle. The site of action of cellular restriction factors are indicated with blue block signs. The site of action of antiretroviral drugs are in white box signs. Reprinted with permission from[65].

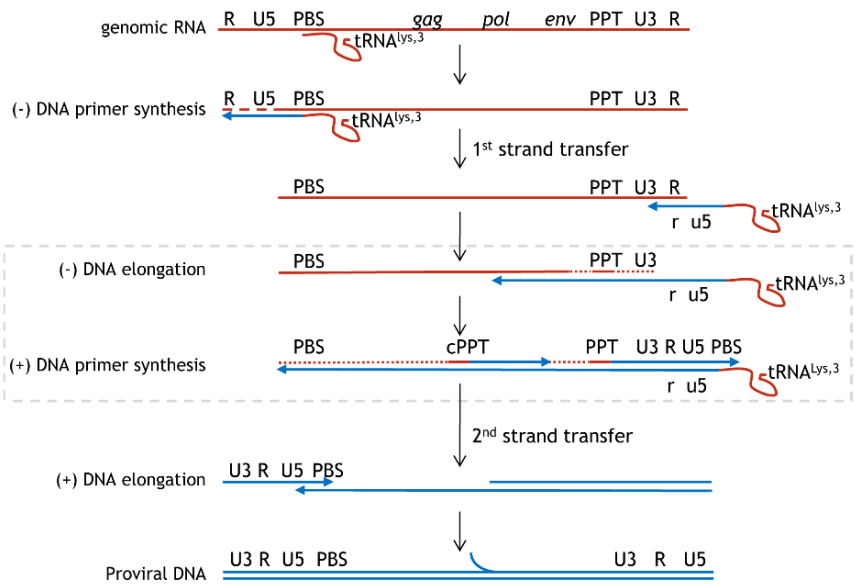


Figure 1.4 HIV reverse transcription. Viral RNA is converted to a double-stranded DNA by reverse transcription. Reverse transcription is a nine-step process. **Step 1**, synthesis of the (-) DNA from the complimentary copy of positive strand of genomic RNA. **Step 2**, removal of the template RNA from RNA-DNA hybrid by the action of RNase H. **Step 3**, first strand transfer. In this step the (–) DNA is transferred from the 5' end to the 3' end of the RNA. **Step 4**, complete synthesis of the full length (–) DNA. **Step 5**, removal of the RNA. The PPT region is resistant to digestive action of RNase H and this residual RNA is later used in synthesis of (+) strand of DNA. **Step 6**, synthesis of (+) DNA is initiated by the PPT RNA primer and extends through the U3-R-U5 LTR. **Step 7**, removal of tRNA and PPT primer by RNase H. **Step 8**, second strand transfer. **Step 9**, extension of both DNA strands by reverse transcriptase. Reprinted with permission from[77].

1.5 HIV Treatment

There has not been any report on a conventional cure for HIV infection. Antiretroviral therapy (ART) on the other hand can suppress replication of the virus in an infected individual, increase peripheral blood CD4+ T cell counts and reverse immunodeficiency[78]. Currently the World Health Organization's aim is to increase the coverage of ART throughout the involved populations using a 90-90-90 plan. The 90–90–90 plan, plans to have 90% of the people living with HIV knowing their HIV status, 90% of the people who know their HIV status receiving ART, and 90% of the people receiving ART having suppressed viral loads. This plan is aimed to be achieved by the end of 2020 and is aligned with the goal of ending the AIDS epidemic by 2030 (Figure 1.5)[79]. Currently global ART coverage for people living with HIV is 59% and 7 out of 10 pregnant women are treated with antiretroviral (ARV) drugs[80]. According to the latest consolidated guidelines released by WHO in 2016, all people living with HIV including children, adolescents and adults, pregnant and breastfeeding women, regardless of clinical stage or CD4+ T cell count (normal 500-1500, immunodeficient <350) should receive lifelong ART soon after diagnosis[81,79]. There are six classes of ARV drugs that target the virus in different steps of replication cycle (Figure 1.3)[82]:

1) Nucleoside reverse transcriptase inhibitors (NRTIs)
2) Non-nucleoside reverse transcriptase inhibitors (NNRTIs)
3) Protease inhibitors (PIs)
4) Integrase inhibitors (INSTIs)
5) Fusion inhibitors (FIs)
6) Chemokine receptor antagonists (CCR5 antagonists)

The most effective method is the so-called highly active antiretroviral therapy (HAART), which is a combination of three or more ARV drugs from different categories. For an adult the most common ART is a fixed dose combination containing the NRTI tenofovir disoproxil fumarate (TDF); the NRTI lamivudine (3TC) or emtricitabine (FTC) and the NNRTI efavirenz (EFV). The patient who is receiving treatment goes under regular viral load monitoring twice in the first year and then every 12 months. Patients who are not stably on ART or are not virally suppressed should start the second and third line of ART. In adults, this regimen consists of two nucleoside reverse-transcriptase inhibitors (NRTIs) plus a protease inhibitor (PI) (e.g. AZT and 3TC as the NRTIs and atazanavir (ATV/r) as the PI)[79]. Besides the use of ART in treatment of infected individuals,

ARV drugs are used before exposure to HIV by people who are not infected but they are at risk. This method of HIV prevention is called pre-exposure prophylaxis (PrEP). TDF alone or in combination with FTC is used commonly as PrEP[79,83]. At risk people are from key population that account for 50% of all new HIV infections worldwide. The key populations include (1) men who have sex with men, (2) people who inject drugs, (3) people in prisons and closed settings, (4) sex workers and (5) transgender people[79,83].

Post-exposure prophylaxis (PEP) is offered to individuals within 72 hours of an exposure that has the potential for HIV transmission. The preferred regimen for an adult is a 28 day prescription of combined TDF, 3TC (or FTC) and lopinavir (LPV/r) or ATV/r[79].

ART is a lifelong treatment that can increase life expectancy of a 20 year old HIV+ patient from 39 years to 70 years[84]. However, numerous side effects have been reported that differ depending on the class of the ARV drug, namely, lipodystrophy (body fat redistribution) by NRTIs due to their abilities to disrupt function of mitochondria and provoke apoptosis[85-87]. For PIs, metabolic abnormalities like hyperglycemia and insulin resistance have been reported[85-87]. NNRTIs are associated with rash and may cause Stevens-Johnson syndrome and toxic epidermal necrolysis [85-87]. Over the last 15 years, as the scale of use of ART in HIV treatment and prevention has increased. Despite the significant benefits and positive contribution to public health of ART, there has been an increase in HIV drug resistance (HIVDR)[80,88]. HIVDR is caused by a mutation or mutations in the genome of HIV that changes the protein that the drug targets so that the ARV is no longer effective[89]. HIVDR in populations happens upon retention in treatment and care services that is caused by serious gaps in ART service delivery, drug stock-outs, negligence to monitor the viral load suppression in the first line of ART and changing the regimen to the second line of treatment, and inadequate support for population adherence to ART[88]. According to WHO's action plan, NNRTI resistance among people on ART who are virally suppressed ranges from 4% to 28%, while it ranged from 47% to 90% among those with unsuppressed viral load on the first-line NNRTI regimens and emphasizes the importance of viral load testing and switching to second-line treatment upon confirmed virological failure[89]. HIVDR has substantial socioeconomical impact. Over the period of 2016–2020, the drug resistance is predicted to be responsible for 135,000 deaths, 150,000 new HIV infections and cost the healthcare system US$ 650 million in sub-Saharan Africa[89]. Addressing all the contributors to HIVDR is of critical importance. A known

contributor to HIVDR is the HIV reverse transcriptase that has no proofreading activity and generates approximately 1 error per genome per replication cycle[90]. A potential contributor to HIVDR is the A3 deoxycytidine deaminases.

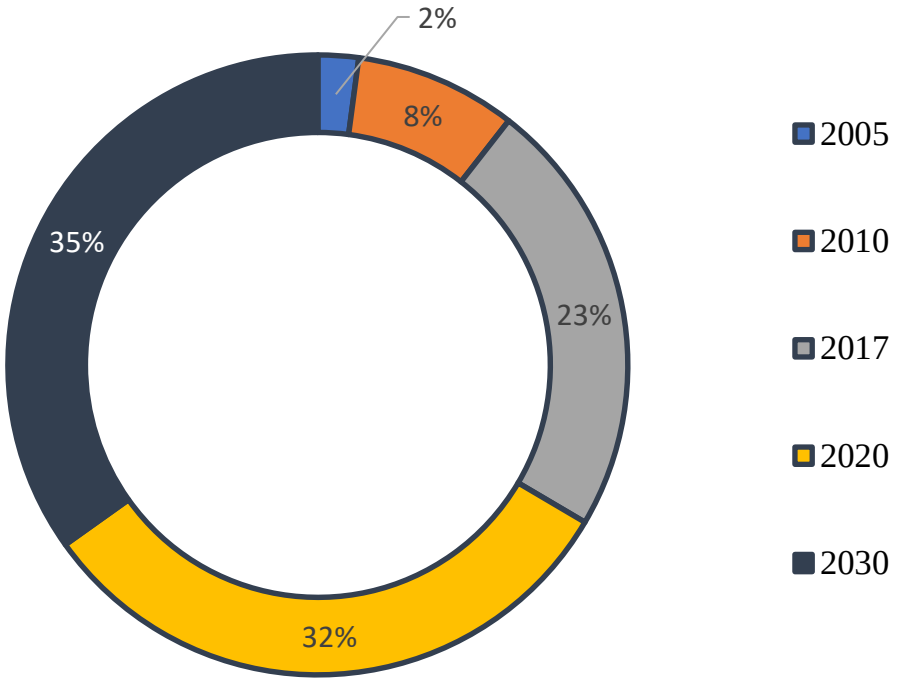


Figure 1.5 Percentage of people receiving ART over time. Adapted from[91].

1.6 APOBEC

The APOBEC family of enzymes are single-stranded cytidine deaminases that are named after the first identified member of the family, Apolipoprotein B mRNA editing complex 1, APOBEC1 (A1)[92]. Other members of this family are activation induced cytosine deaminase (AID), APOBEC2 (A2) , APOBEC3 (A3, there are seven members) and APOBEC4 (A4) (reviewed in ref[77]). These enzymes share similar functional and structural characteristics such as a single or double zinc-coordinating motif which catalyzes the deamination activity[93]. The common amino acid motif of the Zinc coordinating domain is H-X-E-X_{23-28}-P-C-X_{2-4}-C (where X is any amino acid). The histidine (H) and the two cysteines (C) coordinate a zinc atom and form the catalytic pocket. The cytidine will bind to this pocket. The glutamate (E) acts as a proton donor and activates the water molecule that is coordinated by the Zinc atom. The activated water deaminates the ammonium group on carbon 4 of the bound cytidine through a zinc-hydroxide-mediated nucleophilic attack[93]. The Zinc dependent deamination occurs on either mRNA or single-stranded (ss) DNA to form uracil[94-96].

Uracil in DNA is a lesion and the fate of this premutagenic lesion in DNA and subsequently the impact of APOBEC family is determined by the cell type in which the deamination occurs and the cellular DNA repair mechanisms (reviewed in ref [77]) (Figure 1.6). AID is expressed in B cells, and deaminates the ssDNA generated from antibody genes that are undergoing transcription. The resulting uracils act as an initiator for error prone DNA repair to evolve and mature the antibody genes and enable immunoglobulin class switch recombination[97]. A1 is expressed in cells of the human small intestine and is involved in lipid transport. A1 deaminates cytidine at position 6666 of apolipoprotein B mRNA and generates a stop codon (CAA to UAA). This translational stop codon truncates the 550-kD form of apo B, apo B100, to a smaller form, apo B48. Apo B48 is directly involved in lipid transport[98,99]. A2 and A4 do not have any catalytic activity. A2 is expressed in heart and skeletal muscles and plays an important role in muscle development and function[93,100,101]. A4 is expressed in testis tissue and the function of this protein is yet to be determined[102]. A3 enzymes are expressed in T cells, macrophages, and germ cells. Five out of seven members of the A3 family (i.e. A3C, A3D, A3F, A3G and A3H) can suppress replication of HIV-1 in CD4+ T cells in the absence of HIV Vif protein. These members can deaminate cytidines on the nascent cDNA (newly synthetized complementary DNA, (-) DNA) of the retrovirus when

it is single-stranded during reverse transcription. The use of uracils on the (-) DNA as a template for (+) DNA synthesis results in a hypermutated and nonfunctional virus[103] (reviewed in ref[77,92]).

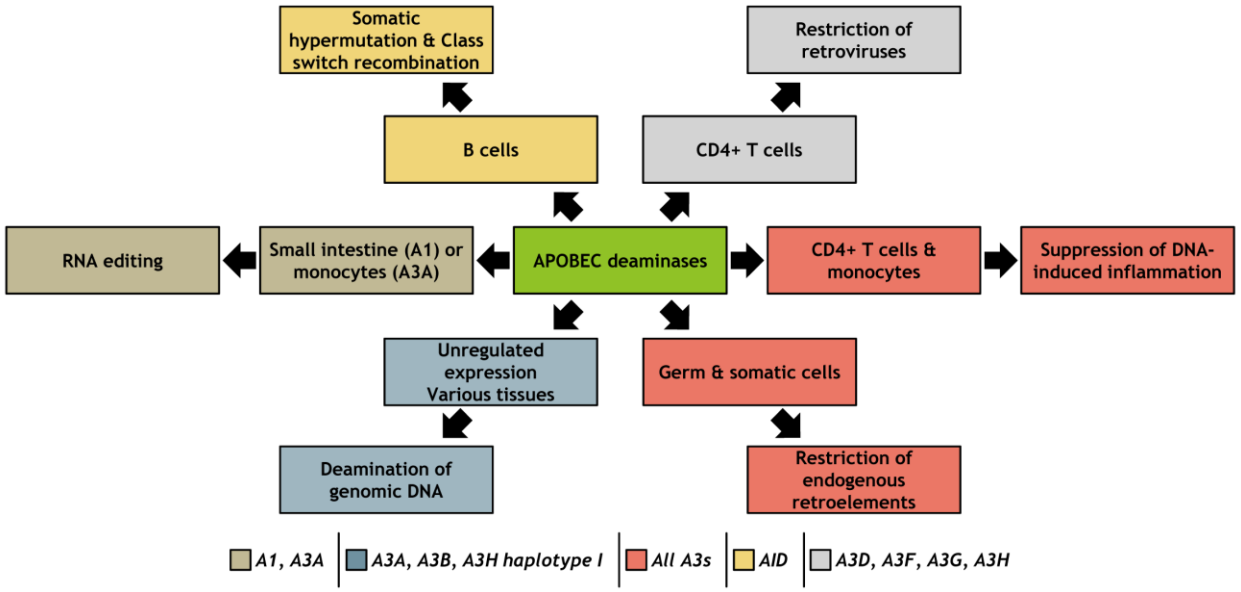


Figure 1.6 Overview of APOBEC family function. APOBEC family members are expressed in distinct cell types. A3A in monocytes and APOBEC1 in the small intestine edit RNA. AID deamination initiate antibody diversification pathways leading to somatic hypermutation and class switch recombination in B-cells. In CD4+ T cells, cytoplasmic A3s cause the deamination of retroviral ssDNA intermediates. A3s in monocytes and CD4+ T cells restrict foreign DNA reducing the DNA-induced inflammatory response. A3s may restrict retrotransposons in germ cells and somatic cells, through RNA binding or deamination of reverse transcripts. Unregulated A3 expression can lead to deaminations in genomic DNA leading to cell transformation or chromosomal instability both of which can lead to cancer. Reprinted with permission from[77].

1.7 APOBEC3

A3 enzymes are evolutionarily conserved cytidine deaminases and are expressed in lymphoid tissues of placental mammals. This family has seven members, A3A, A3B, A3C, A3D, A3F, A3G and A3H [104]. A3G was the first member of the family that was discovered in 2002 by Sheehy *et al*, when they infected two different cell types CEM and CEM-SS with HIV-1 Δ*vif*. They observed that the *vif* deleted virus could not replicate in CEM cells (non-permissive cell line) but the same virus could successfully replicate in CEM-SS cells (permissive cell line). The fact that HIV-1 Vif was crucial for successful replication of the virus in non-permissive cells lead to the discovery of the limiting agent in these cell lines, A3G (originally called CEM-15)[23].

Members of the A3 family share structural and functional characteristics with the rest of the APOBEC family. A3 proteins deaminate cytidine to form uridine in single-stranded DNA and/or RNA in a Zinc dependent manner [92,95,105]. Gene duplication events on chromosome 22 have resulted in two groups of A3 enzymes with family members having either a single zinc (Zn)-coordinating deaminase domain (A3A, A3C and A3H) or double Zn-coordinating domains (A3B, A3D, A3G and A3F) (Figure 1.7). The A3 Zn-coordinating deaminase domain (Z domain) is further classified into three distinct evolutionary clusters Z1, Z2 or Z3[104,106,107]. Z1 and Z2 domains have a Ser-Trp-Ser/Thr-Cys-X_2–4-Cys motif, whereas the motif in Z3 domain is a Thr-Trp-Ser-Cys-X_2-Cys[108]. The A3A, A3C and A3H enzymes with single Z domain have Z1, Z2 and Z3 type domain respectively. The A3 proteins with double Z domain have either one type of Z domain or a combination, for example A3B has (Z1-Z2), A3D (Z2-Z2), A3F (Z2-Z2) and A3G (Z1-Z2) and A3H is the only Z3 type domain[107 104]. However, only the C-terminal Z-domain is catalytically active in these enzymes[109 107].

A3s are considered part of the "intrinsic" immune system and act as host restriction factors against replication of retroviruses and retrotransposons. The mechanism of A3s function has been studied extensively in HIV-1 where, restriction of HIV primarily happens in the absence of Vif. A3C S188I, A3D, A3F, A3G and A3H haplotypes II, V and VII are present in the cytoplasm of infected cells and can restrict replication of HIV-1 in CD4+ T cells in *Δvif* condition (Figure 1.8A). The polymorphism in the population of A3s has different effects on the restrictive properties of the virus. A single nucleotide polymorphism (SNP) in A3C results in a 188I variant that is present at about a 10% frequency in diverse populations throughout Africa and has 10-fold more restrictive

activity than the common 188S A3C. This enhanced antiretroviral activity correlates with ability of this A3C 188S to dimerize. The A3C 188S is a stable monomer, whereas A3C 188I is in equilibrium between monomer and dimer[110]. A3H has seven major haplotypes (hap I to VII) with widely varying HIV restriction activity, due to different stabilities of the protein based on the haplotype. Only Hap II, V and VII are stable in cells and have HIV restriction activity[111,112]. Hap II is the most prevalent of the "active" A3Hs and is present in 50% of Africans/African-Americans[111,113]. Other haplotypes are partially active (Hap I) or not active (Hap III, IV and VI) due to degrees of stability/instability in cells. The instability is due to Δ15N and/or R105G mutation. It has been proposed that A3H evolved to become unstable due to a compromised cost to benefit ratio. Since these enzymes are promutagenic and A3H haplotype I can localize to the nucleus, it has been found to induce mutagenesis of DNA in lung and breast cancer[114] . It has been hypothesized that this "off-target" activity in the genome was beneficial for restriction of an ancient pathogen, but since this pathogen has been fully suppressed the maintenance of A3H comes at a cost to the host[111,115].

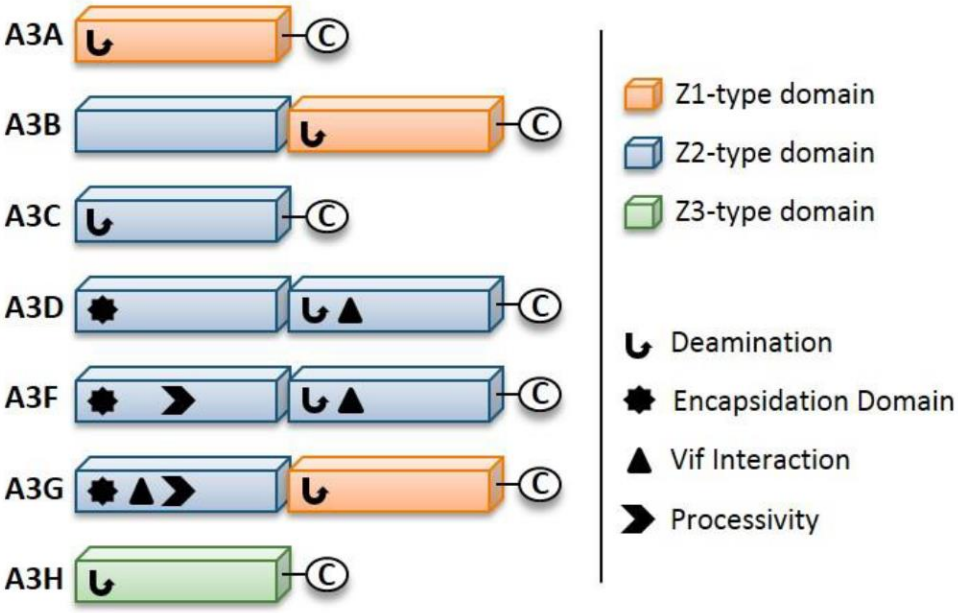


Figure 1.7 Z-type domains of human A3 enzymes. APOBEC3 enzymes contain either a single or double Zn coordinating domain. There are 3 types Z1, Z2, and Z3. A3H is the only Z3 family member. A3B and A3G contain each a single copy of the Z1-type domain. Single domain A3s contain all biochemical characteristics on that single domain whereas double domain enzymes delegate functions to either domain. Reprinted with permission from[92].

1.7.1. Mechanisms of HIV-1 restriction by APOBEC3 enzymes

When a CD4+ T cell is targeted by HIV-1 and the virus fuses with the cell, the capsid of the virus enters into the cytoplasm of the infected cell. The uncoating event, which is the process of removal of the capsid, exposing the viral nucleic acid does not happen immediately. However, the host dNTPs present in the cytoplasm diffuse into the capsid and the reverse transcription process begins[116]. A3s with antiretroviral activity (A3F, A3G, A3H, A3D, and A3C S188I) are present in the cytoplasm of the infected cell but they cannot access the ssDNA of viral genome during reverse transcription due to the timing of the uncoating process and impermeability of the capsid to cellular proteins. For A3s to be restrictive and access ssDNA of the virus during reverse transcription they must become encapsidated into assembling virions in the HIV producer cell by binding to the cellular and viral RNAs. A3s are able to do this by binding to RNAs with the same sequence preferences as the NC domain of Gag that packages the HIV-1 RNA[117]. The encapsidation of A3s is inhibited by the viral protein, Vif. The Vif protein mediates polyubiquitination of A3 proteins by directly binding to the A3 enzyme and recruiting an E3 ligase complex resulting in degradation of the A3s by 26S proteasome (Figure 1.8A)[118]. Due to natural variation in the HIV population there are alleles of Vif with partial activity against A3 enzymes that fail to neutralize one or several A3s[119,120]. Moreover, even a fully active Vif is not able to suppress A3s completely as one to two A3s with potential deamination activity can still get encapsidated in the presence of Vif[25,121-124].

The encapsidated A3s travel with the virus to the next target cell and mediate inhibition of HIV-1 replication by means of two mechanisms deamination-dependent restriction of HIV-1 and deamination-independent restriction of HIV-1.

1.7.2. Deamination-dependent restriction of HIV-1 by APOBEC3s

After entry of the virus to the next target cell, A3F, A3G, A3H, A3D and A3C S188I interfere with HIV replication by catalyzing the formation of numerous uracils from deamination of cytosine in the (-) DNA at “hot spots”. These hot spots are preferred motifs for A3 deaminase activity and are 5'CCC for A3G or 5'CTC for A3H, 5'TTC for A3F and 5'TC for A3D and A3C S188I. The uracils are used as a template during (+) DNA synthesis, which results in C/G→T/A transition mutations that alters the transcription, coding, fitness, replicative capacity and viability

of the provirus after integration into the host genome (reviewed in ref[77]). Alternatively, the uracils may prevent proviral integration. This is because host uracil DNA glycosylase (UNG), a base excision repair (BER) enzyme in the nucleus, can excise the uracils leading to further processing by apurinic-apyrimidinic endonuclease (APE) that cleaves the DNA backbone at abasic sites. If these sites are numerous, this can lead to degradation of the proviral DNA before integration (Figure 1.8B and reviewed in ref[77]).

Although the selected A3s have an overall similar deamination-dependent mechanism to restrict HIV-1, there are some differences in terms of scanning mode, oligomerization and processivity of these enzymes. A3 enzymes recognize cytidines in specific motifs, but they must first find these motifs in a length of otherwise nonspecific ssDNA[125-127]. Scanning is the process of searching the ssDNA for these motifs in an energy independent manner using facilitated diffusion[128,129]. In facilitated diffusion, enzymes use the electrostatic interaction with DNA to facilitate their Brownian motion (Figure 1.9A)[128,129]. There are different modes of scanning, i.e., sliding, jumping and intersegmental transfer[128-130]. Sliding happens when the enzyme diffuses along the phosphate backbone, searching < 20 nt of DNA (Figure 1.9B). In jumping, the enzyme undergoes micro dissociations and reassociations with the same DNA strand without diffusion into the bulk solution (Figure 1.9C). In intersegmental transfer, an enzyme with two DNA-binding sites binds a second site before releasing the first site in a two-step movement (Figure 1.9D). Jumping and intersegmental transfer are efficient for covering > 20 nt of the target DNA but they lack local scanning of small segments[128-130] (reviewed in ref[77,92]). A combination of two or three scanning modes is crucial for a inducing enough deaminations in the limited time that the (-) DNA is single-stranded.

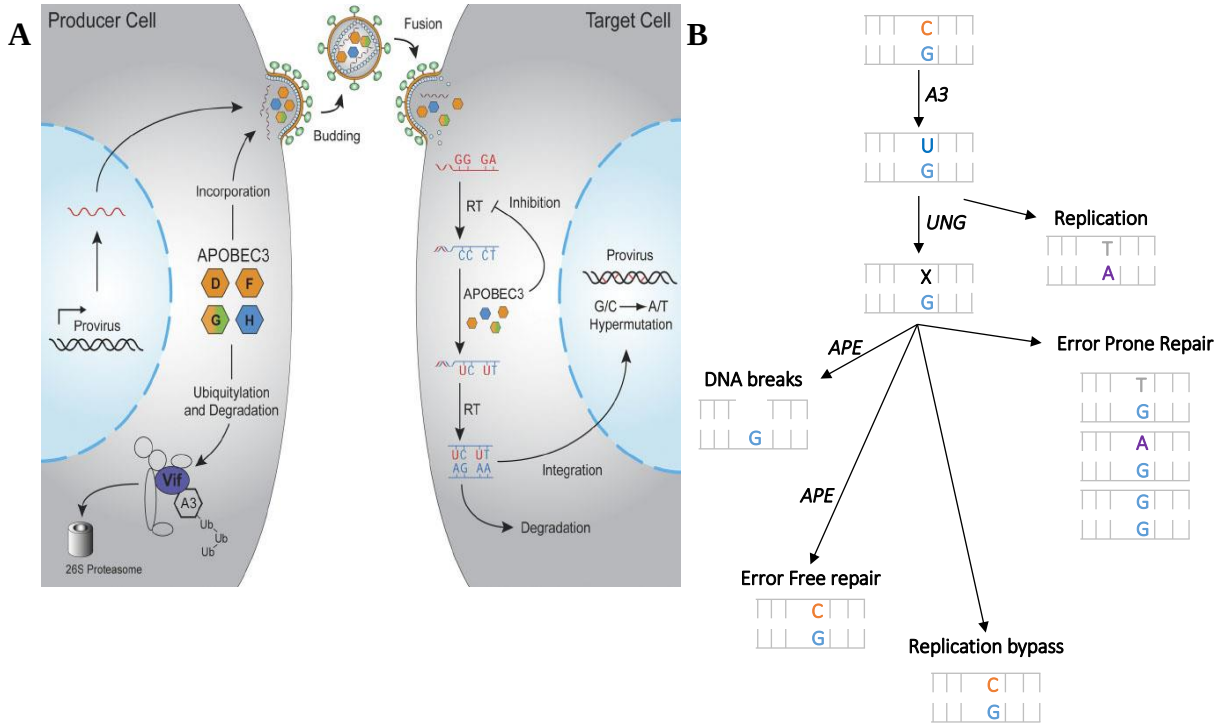


Figure 1.8 APOBEC3-retroviral interaction lifecycle. (A) APOBEC3 mediated inhibition of HIV-1. APOBEC3 proteins are part of the host restriction enzymes against retroelements. These proteins are in the cytoplasm of the infected cells and to be restrictive they should coencapsidate into the assembling viral particles, travel with the particle to the next target cell, then mutate the viral cDNA during reverse transcription. These lesions cause G-to-A hypermutations and change the replicative capacity of the virus. However, HIV Vif mediated ubiquitination and subsequent degradation of A3s by 26S proteasome. Reprinted with permission from[131]. (B) Schematic of the consequences of uracil after cytosine deamination in genomic DNA. In the absence of repair, replication of the U results in a T/A base pair instead of a C/G base pair. Uracils created by cytosine deaminations are removed by uracil DNA glycosylase (UNG) to create an abasic site which can be processed by AP endonuclease (APE). APE cleaves the DNA backbone at the abasic site to enable repair by base excision repair (BER) polymerases or the mismatch repair (MMR) pathway. DNA breaks can occur if the abasic sites are located close together. It is possible for this lesion can be bypassed through error free mechanisms or gap-filling DNA synthesis. Error prone repair can occur by translation synthesis polymerases synthesizing DNA within the gap, depending on cellular conditions. Adapted from[77].

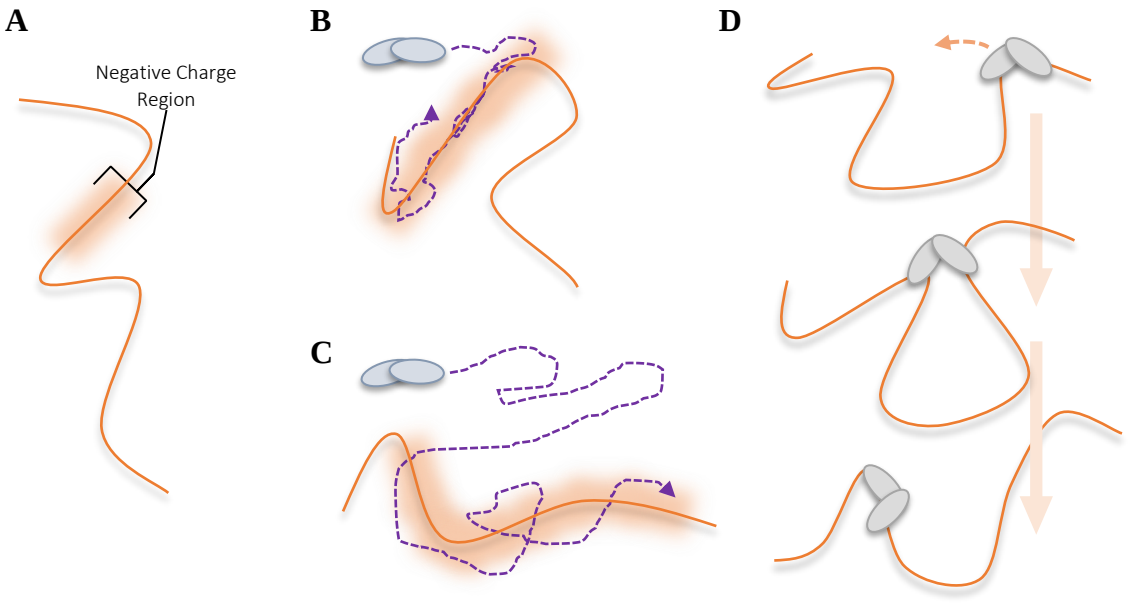


Figure 1.9 APOBEC3 scanning mechanisms. (A) Drawing showing the negative charge region of DNA where A3 enzymes inhabit when they interact with DNA. (B) *Sliding*, short local scanning (<20 nucleotides). (C) *Jumping*, longer distance scans where the A3 continues to remain the negative charge region but can scan over double stranded regions of DNA (>20 nucleotides). (D) *Intersegmental transfer*, this mechanism of scanning involves a doubly bound state. Reprinted with permission from[77] .

1.7.2.1 Processivity of APOBEC3 enzymes involved in restriction of HIV

Processivity is the mechanism of deaminating multiple cytosines in a single enzyme-substrate encounter. The level of processivity differs as a result of A3s' scanning mechanism[132]. A3G enzyme is able to induce deamination using both sliding and intersegmental transfer. This ability benefits the enzyme to have an efficient function[133,134]. A3F is another contributor to restriction of HIV. However, A3F enzyme does so less efficiently than A3G. This can be attributed to lack of sliding ability of A3F. A3F can only scan the target DNA using jumping motions[133]. A3H is the only member of the family with different haplotypes (I-VII), among these only haplotype II, V, and VII can restrict HIV[111]. A3H Hap II and Hap V can utilize both sliding and jumping/intersegmental transfer to facilitate an efficient scanning[135]. The enhanced antiretroviral activity of the A3C S188I polymorphism is due to a change in the dimerization state of the enzyme from monomer to monomer/dimer, with the ability to dimerize being correlated with improved processivity. The dimer enables sliding, jumping, and intersegmental transfer movements of A3C S188I[110]. A3D is the least studied member with antiretroviral activity. This enzyme is less restrictive in comparison to A3F, A3G and A3H and the restrictive capacity of the enzyme is comparable to the less restrictive variant of A3C. However, A3D has not been purified and characterized biochemically to determine if this is due to differences in the processivity. It has been found that A3D in primates is more active and the activity of A3D was lost during evolution from chimpanzees to humans[136,137].

1.7.3 Deamination-independent restriction of HIV-1 by APOBEC3

Besides the deamination dependent restriction of HIV-1 by A3s, which is the main mechanism to inhibit replication of the virus, some A3 enzymes have deamination-independent restriction ability. In deamination-independent restriction, A3s suppress some processes such as tRNA primer binding, initiation and elongation of reverse transcription, plus strand DNA synthesis, or integration[138-143]. One explanation for the mechanism underlying inhibition of these processes is that A3F and A3G bind to the primer template and act as a "road-block" to RT polymerization or interact directly with RT, which slows down the extension of the template[139,144,145]. Deamination independent properties of A3F are stronger than A3G and this mode of restriction accounts for 30% of the total viral restriction by A3F[146]. Whereas, this portion is less than 1% for A3G[146]. The enhanced deamination independent restriction by A3F could be

due to higher affinity of the enzyme in binding to DNA and polymerization of the enzyme on the target DNA that leads to a physical obstacle for the RT[139,144].

1.8 HIV overcomes antiretroviral drugs and host restriction factors

Although HIV only codes for 15 proteins it has a high level of genetic variation that allows the virus to escape immunological and pharmacological barriers. Not only has this genetic variation made it difficult to design a successful vaccine but also drug-resistant HIV strains can be transmitted from one patient to another and leave the infected individual nonresponsive to the medication before it is ever received. To overcome this, it is crucial to understand all contributors to HIV genetic diversity.

A combination of factors contributes to this genetic variation:

1) High replication and mutation rate: As the rest of Retroviridae family, HIV-1 has a high mutation rate that has been evaluated between 10^{-4}-10^{-5} mutations/bp/replication cycle[90] . This high rate of mutation is a result of lack of proof reading ability of reverse transcriptase. This high mutation rate is coupled with high replication rate and production of 10^{11} virions/day/patient and 10^{7}-10^{8} infected CD4+ T cells[147,148].
2) Recombination: HIV is a diploid virus which means every particle carries two RNA strands that come from the same host cell. If a cell gets infected twice (termed superinfection), the host genome will simultaneously harbor two different proviruses. During virion encapsidation one RNA transcript from each provirus can get packaged into a "heterozygous" virion. Upon infection of a new cell with this virion, the reverse transcriptase due to its low processivity can transfer back and forth between the two RNA templates (template switching), which can result in synthesis of a recombinant virus (Figure 1.10A)[149].
3) APOBEC3 selective pressure: Due to natural variation in Vif there are partially active Vif proteins present in the virus population in an HIV+ person. These partially active Vif alleles are less effective in mediating degradation of A3s[150]. Moreover, fully active Vif proteins are not able to fully suppress A3s completely and 1-2 A3s with potential deamination activity can still get encapsidated[151]. The G-to-A mutations induced by A3 proteins have been proposed to modulate HIV-1 fitness and lead into increased genetic variation in the

HIV population by two mechanisms (Figure 1.10B). First, through lethal mutagenesis, dead proviruses are generated that cannot increase genetic variation of the replicating viral population. However, parts of the proviral genome harboring such mutations can get rescued when a replication-competent virus infects the same cell and HIV copackages one faulty and one competent viral RNA genome[152]. This allows for recombination during reverse transcription and leads to presence of sublethally mutated genomes in the replicating viral population. Second, sublethally mutated viruses can be generated directly by low levels of A3-mediated mutagenesis and increase genetic variation of the replicating viral population[152].

The fate of the mutations that are caused by A3s are determined by the location and the frequency of these mutations. It has been shown that if the mutation occurs in TGG codon of tryptophan by A3G or in ATGGA context by A3F, the concomitant G to A mutation changes the TGG codon to a TAG stop codon. Insertion of premature stop codons is considered a lethal mutation and can lead into production of nonviable virions, if not rescued during recombination (Figure 1.11)[77].

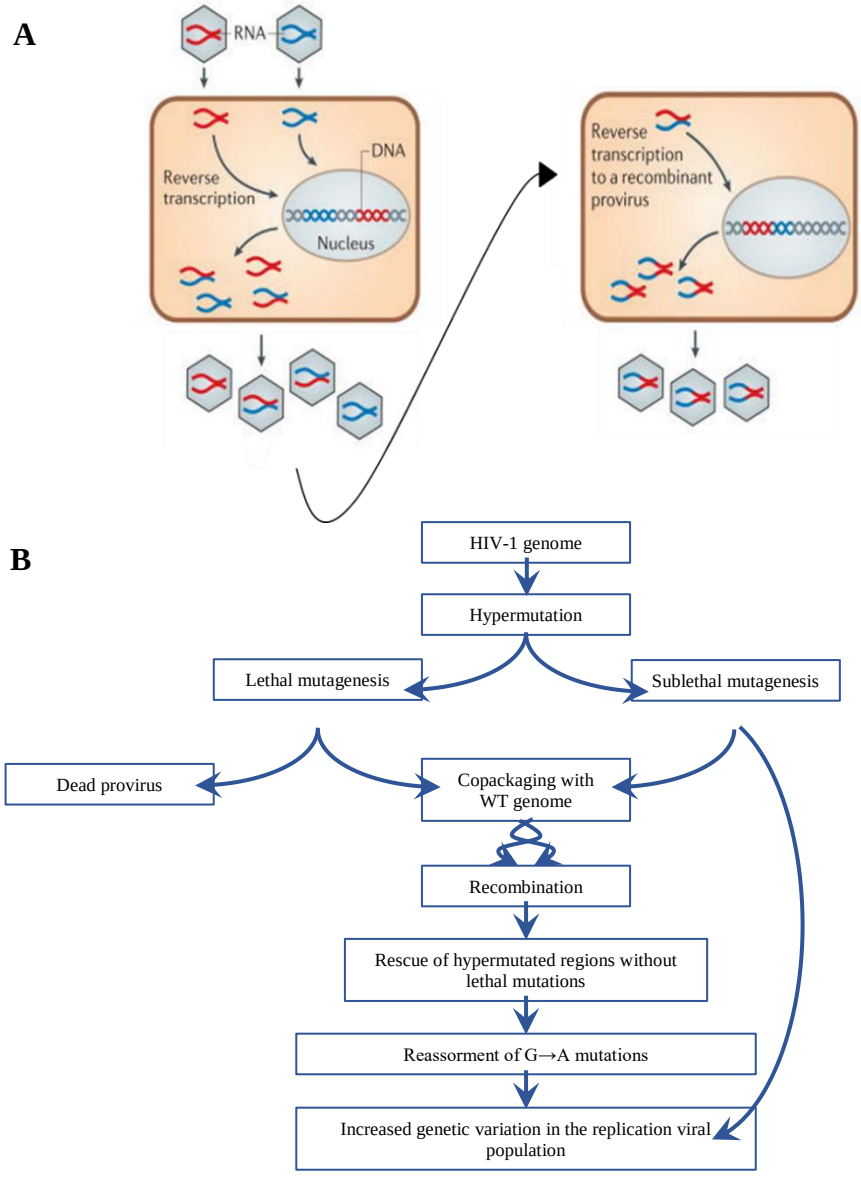


Figure 1.10 APOBEC mutagenesis of HIV. (A) RNA Recombination in HIV. The coinfection with two replication competent virions can cause generation of "heterozygous" virus particles containing RNAs from different parent viruses, if this viral particle infects another cell after template-switch and recombination events, infectious recombinant proviruses are produced and released into the population. Figure adapted from[149]. (B) Flow chart of potential contribution of hypermutation and recombination to HIV genetic diversity. Figure adapted from reference[152].

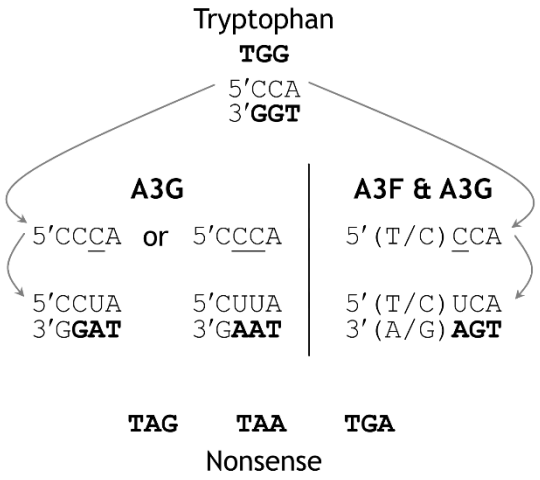


Figure 1.11 Fate of tryptophan codons. The process of insertion of premature stop codons by A3F and AF involving tryptophan in (−) DNA (5′CCA). Reprinted with permission from[77].

1.9 Rationale and Hypothesis

Five of the seven A3 enzymes in humans can restrict the replication of HIV by becoming encapsidated into budding virions and deaminating cytosine to uracil on the (-) ssDNA during reverse transcription. When this (-) DNA is used as a template to synthesize the (+) DNA, the uracil lesions result in transition mutations. The effect of the mutations on the fate of the virion is dependent on the number and position of these mutations. HIV has a genome of ~10^4 bp and has a mutation rate of 1 in 10^4 nucleotides per genome during one replication cycle, which considering the daily turnover of 10^{11} viral particles, results in a viral quasi-species population during chronic infection in the host[147,153,154]. Moreover, the footprint of mutation by A3s in specific sequence contexts in the HIV genome has been found by sequencing integrated proviral genomes from HIV+ people. Analysis of patient proviral sequences has shown ~25% G to A hypermutations in A3 preferred dinucleotide motifs (5′GG→AG for A3G and 5′GA→AA for A3F, A3D and A3H)[155,156]. These data support the fact that even in the presence of Vif, the HIV is unable to completely induce degradation of A3 enzymes such that A3s may contribute to genetic variation of HIV. However, RT errors also have a G→A substitution bias, but in an unidentified surrounding sequence context. Importantly, in the absence of selective pressure, the population of drug resistant mutants is as low as 0.5-1% but they can rapidly grow under drug selection and maintain HIV production[157,158]. Given that, an *in vitro* study by Mulder *et al*. showed that mutagenesis caused by A3s that were not degraded completely by partially defective Vif proteins exceeded the rate of errors generated by RT[150]. They also show that these mutations induced appearance of 3TC resistant variants under drug selection pressure[150]. Conversely, some groups have shown that mutation caused by A3s may lead to a decrease in genetic diversity by reducing the rate of recombination due to decreased homology between the copackaged RNAs[152]. In 2013 Fourati *et al.*, in a study on 30 HIV infected individuals' peripheral blood mononuclear cells (PBMCs) body tissues, and fluids found the presence of drug resistant mutations in PBMCs with no inactivating mutation in *RT* or *PR* sequences and concluded that a combination of cytidine deamination and genomic RNA recombination results in emergence of drug resistant variants[159]. Interestingly, this group showed that the amount of hypermutation in the samples from the tissues exceeded those in PBMC samples and suggested that although these mutations can favor escape of HIV from pharmacological barriers, accumulation of defective viruses may be more dominant in tissues in comparison to PBMCs under pressure of ART[159].

Considering all these findings, we hypothesize that G→A mutations caused by A3 enzymes leads to increased genetic diversity in the replicating population of the virus and that such mutations can lead to drug resistance and escape of the virus from ARV drugs. We further hypothesize that the ability of each member of the A3 family to induce such mutations is variable depending on the polymorphisms.

1.10 Specific Aims

We tested the hypothesis with the following specific aims.

Aim 1: Determine the HIV-1 mutation level caused by A3s in the presence of Vif.

Aim 2: Investigate if A3-induced mutations can contribute to the evolution of the virus and yield in drug resistant variants.

Aim 3: Identify the HIV restriction activity of the A3F 231I and A3F 231V polymorphisms, in cooperation with A3G and alone. We aimed to study the A3F 231I/V polymorphisms because most people carry a heterozygous allele of 231I/231V[136,160].

2.0. COMBINED RESTRICTION OF HIV-1 BY DEOXYCYTIDINE DEAMINASES APOBEC3G AND APOBEC3F UNLIKELY TO PROMOTE EVOLUTION OF DRUG RESISTANCE WITHOUT FUNCTIONAL CONSEQUENCES ON VIRAL REPLICATION FITNESS

Nazanin Mohammadzadeh[1], Robin P. Love[1], Richard Gibson[2], Eric J. Arts[2], Art Poon[2,3], and Linda Chelico[1*]

[1]University of Saskatchewan, Biochemistry, Microbiology, and Immunology, College of Medicine, Saskatoon, Saskatchewan, Canada

[2] Western University, Schulich School of Medicine and Dentistry, Department of Microbiology and Immunology, London, Ontario, Canada

[3] Western University, Schulich School of Medicine and Dentistry, Department of Pathology and Laboratory Medicine, London, Ontario, Canada

*To whom correspondence should be addressed: Linda Chelico, linda.chelico@usask.ca

All experiments and data analysis in this chapter were performed by N.M. Illumina sequencing analysis was conducted collaboratively between A.P., R.P.L. and N.M. L.C., E.J.A., R.G. and N.M. conceived and designed the experiments.

2.1 Abstract

APOBEC3 deoxycytidine deaminases are host restriction factors of HIV-1. The APOBEC3 enzymes that can become encapsidated into HIV-1 virions are able to induce mutagenesis of proviral DNA through the deamination of cytosine in HIV-1 (-) DNA, which forms promutagenic uracil and results in G→A mutations. HIV-1 overcomes this host restriction through the viral protein Vif that binds to and induces APOBEC3 polyubiquitination and proteasomal degradation by acting as the substrate receptor for a host E3 ubiquitin ligase. There is a balance between Vif-mediated degradation of APOBEC3 enzymes and APOBEC3 deaminase activity that determines the number of mutations incurred by a provirus and if the mutations result in viral inactivation. Within this dynamic relationship, the APOBEC3 enzymes have been found to be beneficial, neutral, or detrimental to several aspects of HIV-1 host adaptation such as CTL escape, co-receptor usage and antiviral drug resistance. Here, we assessed the ability of coexpressed APOBEC3F and APOBEC3G to induce HIV-1 drug resistance to several antiviral drugs. We found that the APOBEC3F and APOBEC3G, that has been previously shown to form a hetero-oligomer in cells, enabled partial resistance of APOBEC3F to Vif-mediated degradation with a corresponding increase in APOBEC3F encapsidation and deaminations in the presence of Vif. The resulting deaminations of APOBEC3G and APOBEC3F were above the RT-induced G→A mutation rate and resulted in drug resistant variants. However, the drug resistant variants resulting from APOBEC3-induced mutagenesis were less able to replicate than drug resistant viruses derived from RT-induced mutations alone. Altogether, the data support a model in which APOBEC3 enzymes cooperate to restrict HIV-1, which ensures that viral inactivation is more likely to occur than viral evolution.

2.2 Introduction

In order to infect host cells, the Human Immunodeficiency Virus (HIV) must overcome the activity of several host restriction factors[17,161]. Due to the lifelong infection caused by HIV, there is extensive adaptation between the host restriction factors and their counteracting viral accessory proteins[162-170]. HIV restriction factors can prevent viral budding, decrease nuclear accumulation of viral cDNAs, interfere with reverse transcription and induce mutagenesis of proviral DNA[17,161].While viral evolution to continually evade host restriction occurs constantly through selected mutations resulting from reverse transcriptase (RT) error there may also be a contribution

from a group of host restriction factors[118,171-174]. The APOBEC3 restriction factors of HIV can induce mutagenesis of proviral DNA by deaminating cytosine to form promutagenic uracil in (-) DNA when it is single stranded[17,77,161]. Copying of these uracils during (+) DNA synthesis results in guanine to adenine mutations[17,77,161]. The intensity of these mutations is dependent on both the activity of the APOBEC3 (A3) enzymes and the ability of the HIV protein Vif to suppress A3 activity[92].

The HIV-1 protein Vif mediates the proteasomal degradation of several human A3 enzymes that are expressed in CD4 + T cells[92,175]. HIV-1 Vif interacts with the transcription cofactor CBF-β for stability and also interacts with the Elongin C and Cullin 5 components of a Cullin5 E3 ubiquitin ligase assembly[27,176,177]. Elongin C is an obligate heterodimer with Elongin B and Cullin 5 recruits Rbx2 to the complex[27,176,177]. This enables Vif, which replaces the human SOCS2 protein in this complex to act as the substrate receptor, to bring A3 enzymes in proximity to this complex[178-180]. This results in the polyubiquitination and degradation of A3 enzymes after Rbx2 recruits an E2 that transfers the lysine residues[33,181,182]. Without the action of Vif, higher numbers of A3 enzymes would bind HIV-1 genomic RNA or cellular RNA that is also bound by HIV-1 Gag and be encapsidated into virions[117,183,184]. Of the seven human A3 enzymes, A3C, A3D, A3F, A3G, and A3H interact with Vif, resulting in their proteasomal degradation[185-190]. Even if A3 enzymes escape Vif-mediated degradation, Vif can block encapsidation of A3s into virions by a degradation-independent mechanism and Vif can inhibit the activity of A3G when co-encapsidated[191-194]. Despite the multi-modal action of Vif, A3-mediated mutagenesis of proviral genomes in HIV+ people occurs[156,195-204]. This may be due to fortuitous escape from Vif or that the ability of Vif to induce A3 degradation can decrease over the course of an infection[119,120].

Regardless of how A3 enzymes escape Vif-mediated inhibition, they remain within the capsid of the virus and exert their restriction activity in the next target cell that the virus infects. When the HIV-1 reverse transcriptase synthesizes the (-) DNA from the genomic RNA (gRNA) template, the RNase H domain becomes able to degrade the gRNA, exposing single-stranded regions of the (-) DNA[77,205]. A3 enzymes can only deaminate DNA when it is single stranded (ssDNA) and it has been demonstrated that more A3-induced mutations accumulate in proviral DNA in regions that are the last to become double-stranded. These vulnerable regions are on the 5′-side of both the central and 3′ polypurine tracts (PPTs) and result in two gradients of

mutations[206-208]. The A3 enzymes combat this limited time of substrate availability by acting on ssDNA processively, which greatly increases the deamination efficiency of the enzymes[77,92]. Of the five A3 enzymes that can encapsidate into HIV-1 virions and are targeted by Vif, each have differing activity that can be discerned based on their respective levels of processivity[77,92]. In the absence of Vif, when A3 activity is at its highest, A3G can induce the most mutations and is the most processive[77,92]. As processivity declines so do deamination levels in ΔVif HIV-1 virions where A3H, A3F, and A3C show a decreasing gradient of activity after A3G[77,92,131,133,135,209]. There are also other factors that affect A3 activity, such as the polymorphisms present in the human population[110-113,127,160,210-212]. A3H exists as seven major haplotypes based on a combination of five polymorphisms, but only haplotypes II, V, and VII are stable in cells and can encapsidated into HIV virions and restrict HIV[111,113]. However, depending on the haplotype, only 1% (haplotype VII)[111,113], 5-20% (haplotype V)[111,113], or 6-27% (haplotype II)[111,113,127] of the human population has an A3H genotype that can result in HIV restriction. An A3C polymorphism present in 10% of people of African descent enables A3C dimerization and processivity, resulting in HIV restriction ability [110,209]. Another variable that has been identified to effect A3 restriction activity is the preferred deamination motif. A3G preferentially deaminates in 5′CC motifs on the (-) DNA [125]. This results in 5′GG→5′AG mutations in the (+) DNA. This motif occurs in codons for the amino acids for tryptophan, glycine, aspartate, and glutamate and deamination of cytosines results in the formation of nonsense mutations for tryptophan codons or nonconservative amino acid changes for the other codons [77]. The other A3s all prefer to deaminate 5′TC motifs on the (-) DNA, resulting in 5′GA→5′AA mutations in the (+) DNA[77]. These mutations result in a wide variety of missense mutations and can cause nonsense mutations, but nonsense mutations occur at least 2-fold less than from A3G deaminations, depending on the experimental system[77,213].

Since A3-mediated restriction is based on functional inactivation of proviral DNA, the number and type of mutations that an A3 enzyme induces is an important factor. Low levels of mutations or missense mutations may result in either HIV-1 evolution or have a neutral effect, rather than HIV-1 inactivation[150,152,174,199,214-216]. Studies have demonstrated that due to the A3G preferred deamination motif, A3G-mediated mutagenesis is likely to induce viral inactivation[133,144,213,217-219]. However, the A3s that cause 5′GA→5′AA mutations (A3F, A3H, and A3C) may be more likely to induce HIV-1 genetic diversity and evolution[77,213,219]. These data in combination with data demonstrating that during chronic HIV-1 infection Vif can acquire

mutations that partially decrease its ability to induce A3 degradation suggest that HIV-1 can manipulate A3 activity like a mutational rheostat[119,120,150]. Evidence of this has come from studies that have assessed antiviral drug resistance, CTL epitopes, and co-receptor usage, some of which used less effective Vif variants isolated from HIV+ individuals[150,171,174,214,215,219]. Studies support that A3G and A3F can influence the transition from usage of the CCR5 to CXCR4 coreceptor [171,219]. However, studies addressing how A3 enzymes impact CTL epitopes and antiviral drug resistance have resulted in mixed findings[173,174,214,215,220-222]. For CTL epitopes, data has demonstrated that A3-mediated mutagenesis can either enhance or decrease presentation of peptides on MHC receptors. Although some studies demonstrated that the majority of A3 mutated CTL epitopes result in decreased peptide presentation[174], suggesting A3-mediated immune escape, other studies have demonstrated that the A3 mutated CTL epitopes bias the immune response against functionally inactivated viruses[214]. Thus, there may be a spectrum of effects depending on the A3 enzymes encapsidated and the functionality of Vif.

Similar to data obtained on A3-mediated immune escape, the studies assessing whether A3 mediated mutagenesis can cause drug resistance have led to different conclusions. This may be because each study used a different cell-based experimental system. In Mulder et al., the authors used a less active Vif variant (K22H) isolated from an HIV+ individual to create a hypermutated virus population that was largely unable to replicate, but the integrated proviral genomes contained the M184I mutation in *pol* that confers resistance to the antiretroviral (ARV) drug 3TC (Lamivudine, 2′-deoxy-3′-thiacytidine)[150]. By co-infecting cells with a replication competent virus and the hypermutated ones, the authors were able to show that viral recombination could rescue replication in the presence of high amounts of 3TC[150]. The MT2 cell line used also expressed endogenous A3s, suggesting that A3 enzymes, even over multiple rounds of replication, can result in virus evolution [150]. In contrast, a combined experimental and *in silico* study demonstrated that super-infection of cells and viral recombination between inactivated and competent virus is unlikely to occur and suggested that hypermutated proviruses are a dead-end [152]. Nonetheless, the Vif K22H mutation correlated with increased virological treatment failure in HIV+ individuals, which may be due to increased levels of preexisting drug resistance mutations induced by A3 enzymes, even though many of the proviral genomes may be nonfunctional[120,223,224]. However, none of these studies addressed if there could be drug resistance in a WT HIV-1 infection before A3 hypermutation was induced, which must occur since viral replication is maintained. Another

study addressed this and used CEM-SS cells that do not express A3s and created a CEM-SS cell line stably expressing A3G and showed with WT HIV-1 spreading infections that resistance to 3TC only arose in the cells expressing A3G[216]. This study supports the hypothesis that Vif suppresses A3-mediated hypermutation, but still allows for a small enough number of mutations to occur that provides a benefit to the HIV-1 genetic diversity[216]. However, clinical studies have not found a significant correlation between A3 enzymes and HIV-1 drug resistance[159,225-227]. Namely, studies in which A3 enzymes were found to induce drug resistant mutations concluded that the contribution from A3 enzymes was so minor to the total viral population or that sequences with drug resistant mutations also contained stop codons suggesting that the A3 contribution was insignificant. These observations relate to the principle of purifying selection and that drug resistance mutations induced by A3 enzymes may have too many other collateral mutations to produce replication competent virus or be rescued by viral recombination[152,228].

Overall, both the cell-based and clinical studies have not addressed several factors regarding this topic. First, HIV-1 reverse transcriptase (RT) also has a preference to cause G→A mutations and it has not been established whether the A3 sequence contexts of 5′GG or 5′GA can properly discriminate between A3- and RT- induced mutations[90,148,229]. Second, several studies have used "hypermutated" viruses due to the absence of Vif, use of less active Vif variants, or use of proviral DNA sequences from databases[150,152,172], but the effect of A3s with a wild type (WT) HIV-1 and Vif are not known; in particular, we still do not know the average number of deaminations an A3 can induce in a single round of replication with WT HIV-1. Third, several studies have tested only A3G[150,172,216] even though A3 enzymes are known to be coordinately expressed and A3F and A3G have been shown to interact as a hetero-oligomer that has enhanced HIV-1 restriction activity[144,230,231]. Here we assessed if the A3 enzymes A3F and A3G could cause drug resistance to develop in HIV-1. We chose to study A3F and A3G since they are most commonly the active A3s found in the human population[212]. We examined how the enzymes influenced drug resistance against 3TC and AZT (Azidothymidine, Zidovudine (ZDV)) each alone. Since drug resistance can easily develop with monotherapies, we also tested a mixed ARV regimen of TDF (Tenofovir Disoproxil Fumarate), 3TC, and EFV (Efavirenz) (TLE regimen). We found that the A3F/A3G hetero-oligomer enabled partial resistance of A3F to Vif-mediated degradation with a corresponding increase in A3F encapsidation and deaminations in the presence of Vif. Thus, even though A3F is less active than A3G in ΔVif HIV-1 studies, the combination of

the activities and differential sensitives to Vif leads to almost equal contributions to mutagenesis in WT HIV-1 infections. The resulting deaminations were above the RT-induced G→A mutation rate and resulted in drug resistant variants. However, the drug resistant variants resulting from A3-induced mutagenesis were less able to replicate than drug resistant viruses derived from RT-induced mutations alone. Altogether, the data support a model in which A3 enzymes cooperate to restrict HIV-1, which ensures that viral inactivation is more likely to occur than viral evolution.

2.3 Material and Methods

Spreading infection in PBMCs

Healthy donor PMBCs were isolated from whole blood samples using Ficoll-Paque (GE-Healthcare) density gradient centrifugation and SepMate-50 tubes (STEMCELL Technologies) as per manufacturer's protocol. Cells were cultured in RPMI 1640 media (Hyclone) with 10% heat inactivated FBS (Gibco), and 1% PenStrep (Hyclone). Cells were stimulated for 3 days at 37°C with 10 ng/mL recombinant human IL-2 (Sigma) and 10 μg/mL Phytohemagglutinin-L (PHA-L) (Roche). Cells were then harvested and 8 x 10^5 cells/mL were dispensed to tubes and infected with 0.05 MOI of WT HIV_{LAI}. Cells were gently vortexed and incubated with the virus at 37°C for 16 h. Next, the cells were pelleted at 300 x g for 2 minutes at room temperature. The supernatant was removed and the cells were washed twice with PBS. Cell pellets were resuspended in complete RPMI with IL-2 alone or containing a titration of either: AZT (6.25-50 μM), 3TC (9.38-150 μM) or a 1:2 to 1:32 dilution of the Mixed ARVs TDF, 3TC, and EFV with initial concentration of 100 μM, 300 μM, and 1 μM, respectively. The cells were then dispensed into a well of a 24-well plate. This constituted day 0. Three days later 180 μL of supernatant was harvested and saved at -80°C and 180 μL was replaced to the well with its respective drug-containing media. Samples were taken every three days for the duration of twenty days. On day 21, the supernatant was harvested and applied to freshly stimulated PMBCs, prepared as described. This cycle continued until day 30, when the final supernatant was removed and the cellular genomic DNA was harvested by DNAzol (Invitrogen) treatment, following manufacturer protocol, for provirus sequencing. The frozen virus aliquots were subsequently used for RT assay and viral RNA extraction for cDNA synthesis.

Spreading infection in CEM/CEM-SS

Cells were cultured in RPMI 1640 media (Hyclone) plus 10% heat inactivated FBS (Gibco) medium containing 1% PenStrep (Hyclone). For stimulation, 10 ng/mL human IL-2 (Sigma) and 10 μg/mL PHA-L (Roche) was incubated with cells for 3 days. Cells were then harvested and 8 x 10^5 cells/mL and were dispensed to tubes and infected with 0.25 MOI of WT HIV_{LAI}. The spreading infection then progressed as described for PBMCs. Samples were taken every three days for the duration of thirty days. The final supernatant was removed and the cellular genomic DNA was harvested by DNAzol (Invitrogen) treatment, following manufacturer protocol, for provirus sequencing. The frozen virus aliquots were subsequently used for RT assay and viral RNA extraction for cDNA synthesis.

Spreading infection in U87 CD4+ CXCR4+ cells

Cells were cultured in DMEM (Hyclone) +10% FBS (Gibco). For spreading infections, 1 x 10^5 cells in 1 mL were plated in each well of a 12 -well plate and infected with 0.01 MOI of WT HIV_{LAI} with or without encapsidated A3F/A3G. Then, 16 h post infection, the media was replaced with 1.5 mL of DMEM+10% FBS and 1% PenStrep (Hyclone) alone or containing a dilution of: AZT (6.25-50 μM), 3TC (9.38-150 μM) or 1:2 to 1:32 dilution of the Mixed ARVs TDF, 3TC, and EFV with initial concentration of 100 μM, 300 μM, and 1 μM, respectively. Three days later 180 μL of supernatant was harvested and saved at -80° C and 180 μL of appropriate media was replaced to the well. Samples were taken every three days for the duration of thirty days. The cellular genomic DNA was harvested by DNAzol (Invitrogen), following manufacturer protocol, treatment for provirus sequencing. The frozen virus aliquots were subsequently used for RT assay and viral RNA extraction for cDNA synthesis.

Single cycle infectivity assay

HIV-1_{LAI} Δ*vif* Δ*env* or HIV-1 LAI Δ*env* (500 ng) were cotransfected into 1 x 10^5 293T cells with 25, 50, or 100 ng of A3G in MCS-II of pVIVO2 (Invivogen) vector with a C-terminal HA tag, A3F in MCS-I of pVIVO2 vector with a C-terminal V5 tag, A3F/A3G where A3F was cloned into MSC-I of pVIVO2 with a C-terminal V5 tag and A3G was cloned into the MCS-II of the pVIVO2 vector with a C-terminal HA tag, or Empty pVIVO2 vector. Construction of these vectors has been previously described [144]. GeneJuice (Novagen) transfection reagent was used as

per manufacturer's protocol and cells were maintained in DMEM supplemented with 10% FBS. After 20 h the media was changed and 48 h post transfection the virus-containing supernatants were harvested and filtered through a 0.45 μm polyvinylidene difluoride (PVDF) syringe filter. The cells were washed with PBS, lysed in Laemmli buffer and used for immunoblotting. For the infectivity assay, 1 x 10^4 of HeLa CD4+ CCR5+ LTR *lacZ* cells per well of a 96-well plate were infected with a serial dilution of HIV_{LAI} Δ*vif* Δ*env* or HIV_{LAI} Δ*env* virus in the presence of 8 μg/mL of Polybrene. At 48 h post infection infectivity was measured through a colorimetric detection using a β-galactosidase assay reagent (Pierce). Infectivity of each virus was compared by using the infectivity of the No A3 condition as 100%.

Immunoblotting

Tagged proteins in cells or virions were detected by rabbit anti-HA (1:1000, Sigma) or rabbit anti-V5 (1:500, Sigma). Mouse anti-α tubulin (1:1000, Sigma) and anti-p24 (1:1000, Cat# 3537, NIH AIDS Reagent Program) were used to detect the cell lysate loading control (α-tubulin) and the virus lysate loading control (p24). For cell lysates, 40 μg total protein was used. For virions, a portion of the filtered supernatant was concentrated using Retro-X concentrator (Clonetech) according to manufacturer's instructions and 20 μL was used.

Genotyping PBMCs and CEM cells

The A3F polymorphic sites were amplified using PCR. The presence of the polymorphism that results in S108A was determined using the following primers: (forward) 5′ GTGATATCCCAGCCTGAGCA and (reverse) 5′ CTTCATCGTCCATAATCTTCACG. The presence of the polymorphism that results in I231V was determined using the following primers (forward) 5′ AGCCTATGGTCGGAACGAAA and (reverse) 5′ CTGGTTTCGGAAGACGCC primers. The PCR amplification used Q5 High Fidelity polymerase (New England Biolabs) and cycling conditions were 98° C for 30 sec followed by 30 cycles of 98° C for 10 sec, 64-67° C for 20 sec, 72° C for 20 sec and 2 minutes at 72° C. Amplicons were cloned using the CloneJET PCR cloning kit (ThermoScientific) as outlined in the manufacturer's protocol. Analysis for A3F S108A and A3F I231V polymorphisms included 6-12 clones for each position from PBMCs and CEM cells to determine zygosity.

Vif degradation assay

To compare efficiency of Vif in mediating degradation of A3F, A3G and A3F/A3G, 1 x 10^5 293T cells per well of a 12-well plate were co-transfected with 50 ng of A3F-V5, A3G-HA, A3F-V5/A3G-HA or empty pVIVO2 (No A3) and a titration of Vif_{LAI} expression plasmid (0,10, 25, 50, 100, and 200 ng). GeneJuice (Novagen) used as transfection reagent as described by the manufacturer. DMEM and 10% FBS was used to maintain the cells. Twenty hours post transfection the media was changed and 48 h post transfection cells were washed in PBS and lysed with Laemmli buffer. Vif was detected with HIV-1 Vif monoclonal antibody (1:1000; Cat# 6459, NIH AIDS Reagent Program) and immunoblotting was carried out as described for cell lysates.

RT assay

To measure the viability of the virus we performed a RT assay. The assay relates the functional activity of RT to the level of infection. The assay used 10 µL of supernatant in duplicate. RT buffer (25 mM Tris pH 7.8, 37.5 mM KCl, 2.5 mM $MgCl_2$, 0.25% v/v NP-40, 0.025 U Poly rA p(dT) 12-18, and 1 mM DTT) was used to lyse virions and provide the DNA template. The reaction was incubated for 20 minutes at room temperature for lysis. Then, α-^{32}P-TTP (800Ci/mmol), diluted in RT buffer, was added to each sample and incubated for 2 h at 37°C. Afterwards, the mixture was spotted onto a Whatman DE81 paper (Perkin Elmer) and dried for 10 minutes. This was followed by washing in SSC buffer and 85% ethanol. The paper was then dried for 10 minutes at 65°C and exposed overnight to a phosphoimaging screen. The screen was scanned on a Typhoon Trio Imager and the file was analyzed with Image Quant Software. Purified RT of known concentration was used in the assay to construct a standard curve, which was used to determine the RT concentration in the culture supernatants. The standard curve is shown in Figure 2.1H.

Deep Sequencing

HIV genomic RNA was purified from the supernatant of infected cells using QIAamp Viral RNA Kit (Cat# 52904) as per manufacturer's protocol. The cDNA was synthesized using only the Half RT reverse primer (Table 2.3) and AccuScript High Fidelity Reverse Transcriptase (Agilent). From the virus cDNA, *pol* nucleotides 1792-3536 was amplified by PCR with *Pfu* polymerase (Agilent) using Ext F p7.9 forward primer and Half RT reverse primer (Table 2.3). The product of

this reaction was then amplified (1807-3463 nt) in a nested PCR reaction with *Pfu* using Ext F p7.10 forward primer and Half Pol reverse primer. Primers are listed in Table 2.3. Proviral DNA was obtained by using DNAzol Reagent (Invitrogen) as described in the manufacturer's protocol. The extracted DNA was then treated with DpnI (NEB) for 1 hour at 37°C to digest any contaminating transfected plasmid.

The region to be sequenced for both cDNA and proviral DNA was *pol* nucleotides 2573-3301 nt and was divided into three 360 nt overlapping segments. All segment PCR reactions were performed with Q5 polymerase (NEB) Segment 1 (2573-2933 nt) was amplified using primers D67N_F forward and D67N_R reverse primers. Segment 2 (2789-3149 nt) was amplified using primer E138K_F forward and E138K_R reverse primers. Segment 3 (2941-3301 nt) was amplified using primer M184I_F forward and M184I_R reverse primers. Primers are listed in Table 2.3. For use in the Illumina MiSeq we performed another PCR with the segments cognate MS primer to add the necessary overhang adapter sequences. The amplicons were then sequenced following the 16S Metagenomic Sequencing Library Preparation manual (Illumina) using the reagent described. Prepared samples were sequenced using MiSeq Nano V2 chemistry in a 250x250 paired-end reaction.

Deep Sequencing Analysis

Illumina data was processed through a custom code to remove poor quality reads. The data was formatted to the base substitutions in a dinucleotide context. The data once in this form was analyzed with a custom excel code to determine the level of deamination in an A3 relevant dinucleotide context. Overall deamination frequencies were derived by adding all the in-context deaminations for a sample across all three segments and dividing the sum by the total number of nucleotides sequenced (across all three segments).

2.4 Results

2.4.1 Co-expression of A3G and A3F results in partial protection of A3F from Vif-mediated degradation

In the absence of Vif, A3G and A3F have different abilities to restrict HIV-1. In single-cycle infectivity assays, across a range of transfected A3 plasmid, A3F restricts ΔVif HIV-1 6- to 10- fold less than A3G (Figure 2.1A). However, both A3G and A3F are suppressed by Vif in a

single-cycle infectivity with WT HIV-1, as demonstrated by high HIV-1 infectivity levels (Figure 2.1B). To more accurately assess the impact of A3G and A3F, which are coordinately expressed in CD4+ T cells and hetero-oligomerize, we used a vector with two multiple cloning sites so that on a per cell basis each transfected cell would express both A3G-HA and A3F-V5. Under these conditions, the ΔVif HIV-1 restriction by hetero-oligomerized A3F and A3G was equal to that of A3G alone (Figure 2.1A) and the WT HIV-1 appeared to be able to suppress both A3G and A3F (Figure 2.1B). To track the A3G and A3F we immunoblotted cell and virus lysates from the infectivity experiment. The A3G-HA was sensitive to Vif-mediated degradation, resulting in less A3G encapsidation in the presence of Vif, although A3G was still present in cells infected with WT HIV-1 (Figure 2.1C). In contrast, the A3F-V5 was not detected in cells or viruses in the presence of Vif (Figure 2.1D). Upon co-expression of A3G and A3F, we found no change in A3G sensitivity to Vif-mediated degradation or encapsidation, but there was a change in A3F protein levels (Figure 2.1C-D). In ΔVif HIV-1 infected cells, the co-expression of A3G with A3F stabilized the steady state level of A3F in cells and resulted in 2-fold more A3F being encapsidated (Figure 2.1D). During a WT HIV-1 infection, the presence of A3G resulted in incomplete degradation of A3F and resulted in A3F encapsidation (Figure 2.1D). Despite the hetero-oligomerized A3s being encapsidated in the presence of Vif, there was no significant effect on viral infectivity (Figure 2.1B). For A3G, this is likely due to degradation-independent Vif-mediated inhibition of A3G deamination activity. Vif binds A3G on N-terminal domain (NTD) loop 7 that is also required for A3G processivity, resulting in less efficient deamination during reverse transcription[193,194]. This is unlikely to occur for A3F since Vif binds A3F in the C-terminal domain (CTD), which is not required for A3F processivity. However, A3F is also less active than A3G, making the contribution of A3F encapsidated during a WT HIV-1 infection unclear. To determine the extent that A3F is protected from Vif-mediated degradation when co-expressed with A3G, we co-expressed the A3s with different amounts of Vif expressed from a plasmid, rather than the HIV-1 genome. We determined that A3F was more sensitive to Vif-mediated degradation than A3G (Figure 2.1E-F). When co-expressed with A3G, the A3F was not sensitive to Vif-mediated degradation whereas A3G degradation was the same as when A3G was expressed alone (Figure 2.1G). Altogether, these results demonstrate that there is not complete Vif-mediated degradation of A3F when it is co-expressed with A3G. These data prompted us to determine the

A3-induced mutation frequency of HIV-1 proviral DNA in the presence of Vif, since it may have been previously underestimated due to the use of single A3s.

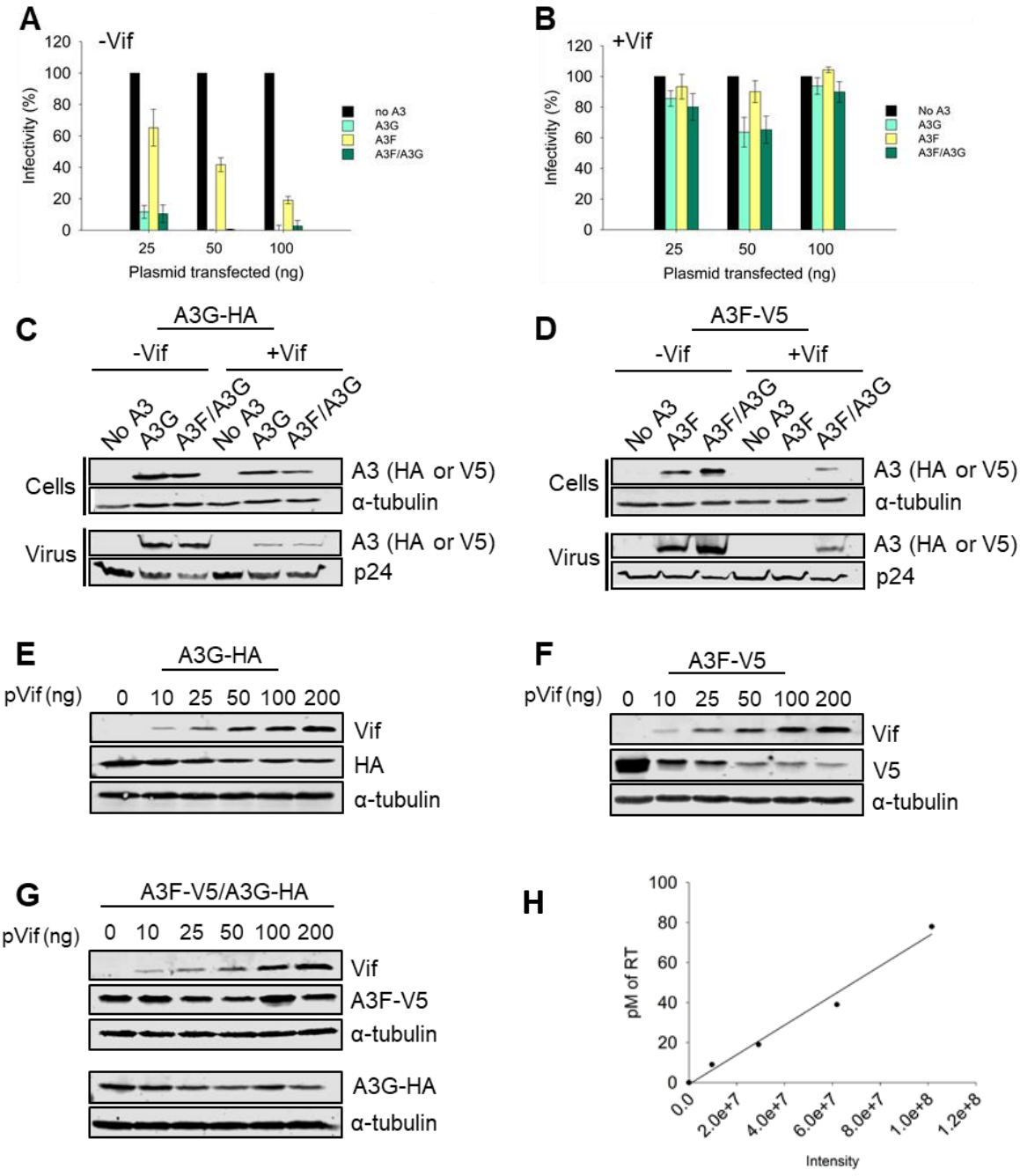

A
-Vif
Infectivity (%)
120
100
80
60
40
20
0
25
50
100
Plasmid transfected (ng)
no A3
A3G
A3F
A3F/A3G
B
+Vif
Infectivity (%)
Plasmid transfected (ng)
No A3
A3G
A3F
A3F/A3G
C
A3G-HA
-Vif
+Vif
No A3
A3G
A3F/A3G
No A3
A3G
A3F/A3G
Cells
A3 (HA or V5)
α-tubulin
Virus
A3 (HA or V5)
p24
D
A3F-V5
-Vif
+Vif
No A3
A3F
A3F/A3G
No A3
A3F
A3F/A3G
Cells
A3 (HA or V5)
α-tubulin
Virus
A3 (HA or V5)
p24
E
A3G-HA
pVif (ng)
0 10 25 50 100 200
Vif
HA
α-tubulin
F
A3F-V5
pVif (ng)
0 10 25 50 100 200
Vif
V5
α-tubulin
G
A3F-V5/A3G-HA
pVif (ng)
0 10 25 50 100 200
Vif
A3F-V5
α-tubulin
A3G-HA
α-tubulin
H
pM of RT
100
80
60
40
20
0
0.0
2.0e+7
4.0e+7
6.0e+7
8.0e+7
1.0e+8
1.2e+8
Intensity

Figure 2.1 Coexpression of A3G and A3F results protection of A3F from Vif mediated degradation. (A) In a ΔVif HIV-1 single-cycle infectivity experiment, A3G and coexpressed A3F and A3G (A3F/A3G) restrict viral replication more than A3F alone. (B) In a HIV-1 single-cycle infectivity experiment A3G, A3F, and A3F/A3G are suppressed by Vif. (A-B) Data were produced from three independent experiments and error bars represent the standard deviation from the mean. (C-D) Immunoblots from the samples used in panels A and B. (C) Vif induces degradation of HA-tagged A3G when expressed alone or coexpressed with V5-tagged A3F. (D) Vif induces degradation of V5-tagged A3F, but A3F is partially rescued from degradation when expressed in the presence of HA-tagged A3G. (E-G) Further analysis of the A3G-mediated protection of A3F from Vif was carried out using increasing amounts of Vif expression plasmid that was cotransfected with a constant amount of (E) HA-tagged A3G, (F) V5-tagged A3F, and (G) co-expression of the HA-tagged A3G and V5-tagged A3F (A3F/A3G). The data show that A3F, but not A3G is protected from Vif mediated degradation when A3F and A3G are coexpressed. (E-G) Immunoblots were conducted from three independent experiments and a representative blot is shown.

2.4.2 Level of A3G and A3F induced G→A mutations exceed those of RT in a single cycle HIV-1 infection

To determine the maximum numbers of A3-induced mutations during a WT HIV-1 infection we amplified the *pol* (nt 2110-2862) from proviral DNA of single-cycle infectivity experiments using 25 ng transfected A3 plasmid and sequenced the amplicons on an Illumina MiSeq platform (Figure 2.2A). The sequences were processed through a custom code to determine the flanking nucleotides of deaminations (see Materials and Methods). The proviral DNA was analyzed for the G→A mutation frequency in the presence and absence of Vif. In the absence of A3 enzymes and in the presence or absence of Vif, the RT rate of G→A mutations stayed relatively the same and was ~0.5 mutations/kb (Figure 2.2B). In the absence of Vif, both A3F and A3G induced high numbers of mutations with A3F inducing 5.6 mutations/kb and A3G inducing 14 mutations/kb (Figure 2.2B). From these values, we defined A3-induced G→A hypermutation as being at least 10-fold above the RT G→A mutation rate. In the presence of Vif, the A3F-induced mutations were not significantly different from the no A3 condition, consistent with the immunoblot that showed no detectable A3F encapsidation in the presence of Vif (Figure 2.2B, 0.5 mutations/kb). In the presence of Vif, A3G was able to induce 2.1 mutations/kb, although this appeared not to have any significant effect on the infectivity level of the HIV-1 (Figure 2.2B and Figure 2.1B). When A3F and A3G were coexpressed in the absence and presence of Vif, the mutation frequency was similar to A3G alone and was 11 mutations/kb and 1.8 mutations/kb, respectively (Figure 2.2B). These data indicate that in a WT HIV-1 infection, the A3s can induce a mutation frequency in proviral DNA that is ~4-fold above RT in a single-cycle of replication and suggest that hypermutated proviral DNA isolated from HIV+ people was exposed to A3s over multiple rounds of infection.

To determine which A3 was inducing the G→A mutations, we further analyzed the sequencing data to determine the GG→AG and GA→AA mutation frequencies. Since A3G deaminates within 5′CC contexts, the (+) DNA mutations are GG→AG. This are discernable from A3F deaminations that occur within 5′TC contexts that result in (+) DNA mutations that are in the GA→AA context. Analysis of the ΔVif HIV-1 condition demonstrates that the random RT mutations that fall into an A3 context are equivalent to the total RT-induced G→A mutations, demonstrating that in context G→A mutations *a priori* cannot solely be attributed to A3 deamination activity (Figure 2.2B). The A3F-induced mutations are 91% in the GA→AA context

(Figure 2.2B). The A3G-induced mutations are less stringent with the preferred GG→AG context being mutated 79% of the total (Figure 2.2B). In the WT HIV-1 condition when each A3 is expressed alone the A3F-induced mutations are 69% in the GA→AA context and the A3G-induced mutations are 87% in the GG→AG context (Figure 2.2B). In the ΔVif HIV-1 condition with combined A3F/A3G expression, the GA→AA mutations and GG→AG mutations were similar to A3G alone (Figure 2.2B, A3F/A3G has 2.8 GA→AA and 8.5 GG→AG mutations/kb), likely due to high A3G activity that can obscure the cooperative effects of coencapsidated A3F/A3G [144]. In the WT HIV-1 condition, the mutations induced by the combined expression of A3F/A3G results in a 60/40% split between GG→AG/ GA→AA mutations with the total amount of GA→AA mutations being more than when A3F or A3G were expressed alone (Figure 2.2B). This was consistent with Vif-mediated degradation data that showed A3F can be partially rescued from Vif mediated degradation in the presence of A3G (Figure 2.1E). Altogether, the data demonstrated that the combined expression of A3G and A3F resulted in a combined effect of both A3s on proviral mutagenesis in the presence of Vif, likely due to the ability of A3G and A3F to form a hetero-oligomer[144].

2.4.3 A3F and A3G can induce drug resistance mutations in proviral DNA

We next determined the nature of the mutations in the proviral DNA to establish whether A3F and A3G could contribute to preexisting drug resistance mutations in proviral DNA. We examined several drug resistance changes along the *pol* gene and noted a significant increase in mutations from 5′→3′ in the (+) DNA, consistent with gradients of mutations forming in proximity to polypurine tract (PPT) regions[206-208]. The HIV-1 central PPT (cPPT) is within *pol*, with mutations peaking directly before the cPPT and within *pol*. We examined two types of possible mutations, those that contain an A3 deamination motif in the codon, which correspond to drug resistant mutations D67N, M184I, E138K, and G190R/E (Table 2.1). We also determined the level mutations that can only be induced by RT, such as K65R, T215Y/F, K103S/N, and Y118L/C/H to determine if the ability of A3s to interfere with RT polymerization also effects its insertion fidelity (Table 2.1). We used the 'No A3' transfection condition to represent the RT induced mutations. In the absence of Vif, A3F was able to induce the D67N mutation 2-fold, the M184I mutation 7-fold, the E138K mutation 11-fold, the G190R 6-fold, and the G190E 15-fold more than RT. In contrast, in the absence of Vif, A3G preferentially caused mutations more than RT at E138K (23-fold more

than RT), M184I (108-fold more than RT), G190R (74-fold more than RT), and G190E (2-fold more than RT), but the D67N mutations were not significantly different from the no A3 condition (Figure 2.2C-F). When the A3F and A3G were coexpressed in the absence of Vif, the mutation pattern matched both A3F and A3G alone indicating that the hetero-oligomerization and coencapsidation of A3F and A3G did not alter their deamination preferences (Figure 2.2C-F). For A3F in the presence of Vif, only the D67N mutations (3.5-fold) and G190E (2-fold) were significantly mutated in comparison to the No A3 condition (Figure 2.2C-F). For A3G in the presence of Vif, there were more mutations induced than for the No A3 condition at the M184I site (3-fold), E138K site (1.3-fold), and G190R (2.8-fold) (Figure 2.2C-F). When A3F and A3G were coexpressed, only the mutations at the D67N site (1.2-fold more than RT), M184I site (2-fold more than RT), G190R site (1.6-fold more than RT) and G190E site (9.5-fold more than RT) were preferentially mutated (Figure 2.2C-F). Altogether, these data suggest that in a WT HIV-1 infection, A3G preferentially creates the M184I and G190R mutations and A3F preferentially creates the D67N and G190E mutations.

We also analyzed if drug resistant mutations that can only be induced by RT error, since they do not have cytosines in the (-) DNA sequence of the codon, were present in proviral DNA. We found that RT could induce mutations at K65R, K103S/N, Y188C/H/L, and T215F/Y, among others (Figure 2.2G-J. For the K103S/N, Y188C/H/L, and T215F/Y mutations, the No A3 condition in the absence or presence of Vif had similar levels of these mutations (Figure 2.2G-J). However, for the K65R, the presence of Vif resulted in a 14-fold increase in the frequency, but overall this mutation occurred at a low level (Figure 2.2F, $4x10^{-5}$, -Vif; $6x10^{-4}$, +Vif). In Δ Vif condition the K65R mutation in the presence of A3F, A3G or A3F/A3G had 6- to 8- fold increase (Figure 2.2G). This effect on the K65R mutation was not observed in the presence of Vif (Figure 2.2G). For the K103S/N mutation, the presence of A3G, either in the absence or presence of Vif resulted in an absence of the K103S mutation, although the K103N mutation was not changed (Figure 2.2H). For the Y188C/H/L mutation, the presence of A3G in the absence of Vif resulted in 5-fold less Y188C mutations (Figure 2.2I). However, in the presence of Vif, A3F, A3G, and A3F/G all caused changes in the mutation frequencies. The presence of A3F or A3G decreased the amount of Y188L mutations 3-fold (+Vif, A3F) to 7.5-fold (+Vif, A3G) and increased the amount of Y188H mutations by 1.8-fold (+Vif, A3F and A3G) (Figure 2.2I). The A3F/A3G condition resulted in 2-fold less Y188C mutations (Figure 2.2H). In the presence of Vif, for the T215F/Y

mutation, the presence of A3F resulted in a 2-fold increase in the mutation rate (Figure 2.2J). Since the (-) DNA sequence for these codons have no cytosines, these data suggest that either the physical inhibition of RT by A3 enzymes or nearby A3-catalyzed uracils can influence the RT insertion fidelity. However, these experiments were done in the absence of drug selection, so we sought to test which mutations were selected for in the presence of drug.

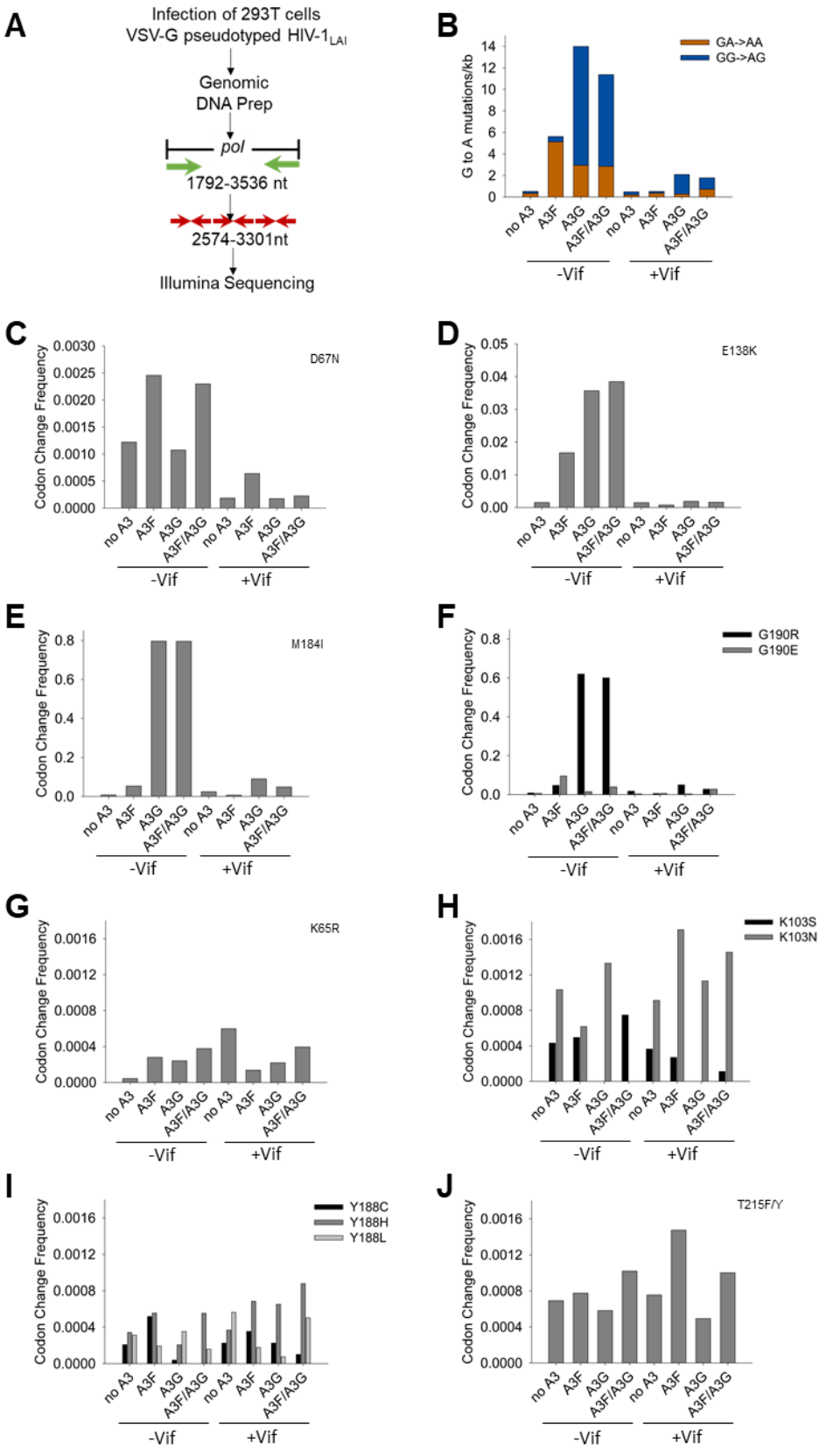

A
Infection of 293T cells
VSV-G pseudotyped HIV-1LAI
Genomic
DNA Prep
pol
1792-3536 nt
2574-3301nt
Illumina Sequencing
B
G to A mutations/kb
GA->AA
GG->AG
no A3
A3F
A3G
A3F/A3G
-Vif
+Vif
C
D67N
Codon Change Frequency
D
E138K
E
M184I
F
G190R
G190E
G
K65R
H
K103S
K103N
I
Y188C
Y188H
Y188L
J
T215F/Y

Figure 2.2 Drug resistant mutations can be induced by A3-mediated mutations and RT-induced error in a single round of virus replication. (A) Proviral DNA resulting from single cycle infectivity assays from Figure 2.1 was used as a template to PCR amplify 1011 nt of *pol* (2085-3096), which was further amplified with nested primers that contained Illumina adapters, resulting in a total 762 nt region of *pol* that was analyzed by next-generation sequencing. The 762 nt was amplified as three separate segments. (B) The G→A mutations in both the GA→AA (A3F context) and GG→AG (A3G context) were determined from ΔVif and WT HIV-1 infections. (C-E) Drug resistant mutations that can be induced by A3-mediated mutagenesis due to a deamination motif within the (-) DNA sequence. Note that not all y-axis scales are the same due to different mutation frequencies imposed by the timing of (+) DNA synthesis. (C) The D67N mutation was preferentially induced by A3F. (D) The E138K mutation was induced by A3F and A3G in ΔVif, but not WT HIV-1 infections. (E) The M184I mutation was preferentially induced by A3G. (F-I) Drug resistance mutations that can only result from RT-induced error. The RT-indcued errors resulted mutations at (F) K65R, (G) K103S/N, (H) Y188C/H/L, and (I) T215F/Y. Drug resistant mutations relevant to the drug resistance selection experiments are shown.

2.4.4 Drug resistant HIV-1 can be selected for from spreading infections of PBMCs

The A3G and A3F could contribute to drug resistance mutations in the absence of Vif and any drug selection, but with mutation rates of 4 mutations/kb even in the presence of Vif there would be a chance that the A3s would induce enough mutations to inactivate the virus. Despite this, there is evidence that inactivated drug resistant viruses can be rescued, if they are transcriptionally active, through viral recombination of two distinct genomic RNAs[150]. As a result, we determined if these mutations could cause viral growth in the presence of ARVs with endogenously expressed A3s from PBMCs. A3F commonly occurs as two polymorphisms, 108S/231I and 108A/231V[160,212]. The donor genotype for A3F was 108S/231I, which was the same A3F in the molecular clone used for single-cycle infectivity assays (Figure 2.1 and Figure 2.2). The selection of drug resistant virus used a two-fold serial dilution of the ARV, either AZT (6.25-50 μM), 3TC (9.38-150 μM), or the Mixed ARVs (TDF, 3TC, and EFV; TLE regimen) (Figure 2.3A). Mutations conferring resistance to AZT (D67N) and 3TC (M184I) can directly arise from A3-mediated mutagenesis (Figure 2.2C, D and Table 2.1). In the Mixed ARVs, the 3TC resistance (M184I), EFV resistance (primarily K103S/N, Y188L/C/H, and G190E/R), and TDF resistance (K65R) can arise from a combination of RT- and A3- mediated mutagenesis (Figure 2.2D-H). For A3-mediated mutations, only the M184I and G190E/R mutations are possible (Table 2.1). After infection of PBMCs with WT HIV-1_{LAI}, the RT activity in supernatants was measured every third day and used to detect if virus was present. The growth took place over 30 days with the total viral supernatant being added to fresh PBMCs from the same donor at day 21. The normal growth of HIV-1 in the PBMCs resulted in a peak at day 12 (Figure 2.3B). Proviral DNA was isolated for sequencing at day 30 and genomic RNA was isolated for cDNA synthesis and sequencing at the peak of virus replication.

In the absence of any ARV drugs, the PBMCs accumulated drug resistance mutations. The highest levels of mutations in the proviral DNA were at K103N and G190E, conferring resistance to EFV (Figure 2.3C). The gRNA had corresponding K103N and G190E mutations, although the G190E was significantly lower in gRNA consistent with the principle of purifying selection since G190E has been reported to create an aberrant protease cleavage site[232]. Mutations at other sites such as M184I and D67N that can be induced by A3G and A3F, respectively, were also present, consistent with the single-cycle infectivity data showing the preexistence of drug resistance

mutations (Figure 2.3C, Figure 2.2E, G). When the HIV-1 was grown in the presence of ARVs, growth curves with delayed peaks were detected, but with approximately 10-fold less virus than in the absence of drug, consistent with low level variants emerging from the culture (Figure 2.3B and Figure 2.3D, F, H).

Each ARV selected for a unique combination of drug resistance mutations. The HIV-1 resistance to 3TC occurred at 75 μM 3TC and correlated with high levels of the M184I mutation in gRNA (Figure 2.3D-E). The G190E mutation was much higher, despite not being required for 3TC resistance, but suggests that G190 was a hot spot for mutation either by RT or A3s. The HIV-1 resistance to AZT occurred at 12.5 μM AZT and resulted in a similar growth curve, but with a more delayed peak of RT activity compared to growth with 3TC (Figure 2.3F, day 21 compared to day 15). The high level of resistance to AZT conferred by the T215F/Y (RT induced mutation) and D67N (A3F induced mutation) was observed at low levels in culture likely due to the selection of drug resistant variants at a low concentration of AZT (Figure 2.3E). Consistent with the M184I mutation causing increased sensitivity to AZT, the mutation was found in the proviral DNA at a much greater extent than in the gRNA (Figure 2.3E). The Mixed ARVs should select for resistance mutations at K65R, M184I, and any of the following K103S/N, Y188L/C/H, and G190E/R. At the 1:2 dilution, we recovered resistant virus that had mutations at K65R, M184I, and K103S (Figure 2.3H-I). The G190R was present at high amounts in proviral DNA but was 3-fold less likely to be in gRNA compared to the K103S (Figure 2.3I and inset). Consistent with M184I conferring increased sensitivity to TDF, there was a much higher amount of M184I mutations in the proviral DNA, than the gRNA (Figure 2.3I and inset). Altogether, the data from virus growth on infected PBMCs with and without drug selection demonstrated that this experimental system could be used to isolate drug resistant viruses and recover drug resistant mutations through deep sequencing. However, the data did not enable us to interpret the contribution of A3 enzymes to the emergence of drug resistant viruses or if A3s could induce the emergence of replication competent drug resistant viruses in a single round of replication.

Table 2.1 HIV-1 nucleotide and non-nucleotide reverse transcriptase inhibitors and possible resistance mutations.

Drug Class	Drug	Primary Drug Resistance Mutations[1]	Possible A3-induced Mutation	Codon in (+)/ (-) DNA	Probable APOBEC3[2]
NRTI	(3TC) Lamivudine	*K65R* **M184I/V**	M184I	5′ATG/3′CAT	A3F or A3G
	Abacavir Didanosine	**K65R** *K70E* **L74V/I** **Y115F** M184I/V	M184I	5′ATG/3′CAT	A3F or A3G
	Tenofovir (TDF) Stavudine	M41L **K65R** D67N *K70E/R* L210W **T215Y/F**	D67N	5′GAC/3′CTG	A3F
	Zidovudine (AZT)	M41L D67N K70R L210W **T215Y/F**	D67N	5′GAC/3′CTG	A3F
NNRTI	Etravirine Rilpivirine	L100I K101E/P E138A/G/K/Q Y181C/I/V Y188L G190E M230L	E138K G190E	5′GAA/3′TTC	A3F
	Nevirapine Efavirenz	L100I K101E/P K103N/S V106A/M Y181C/I/V Y188L/C/H G190A/S/E M230L	G190E	5′GGA/3′TCC	A3G

[1]Drug resistant mutations are listed in bold face type if they are associated with the highest levels of resistance, mutations in italics contribute to resistance, and mutations in plain type contribute to resistance in combination with other resistance mutations. Data from the Stanford University HIV Drug Resistance Database.

[2]The probable APOBEC3 is based on the preferred sequence context for deaminase activity in the viral (-) DNA which is 5′TC for A3F and 5′CC for A3G.

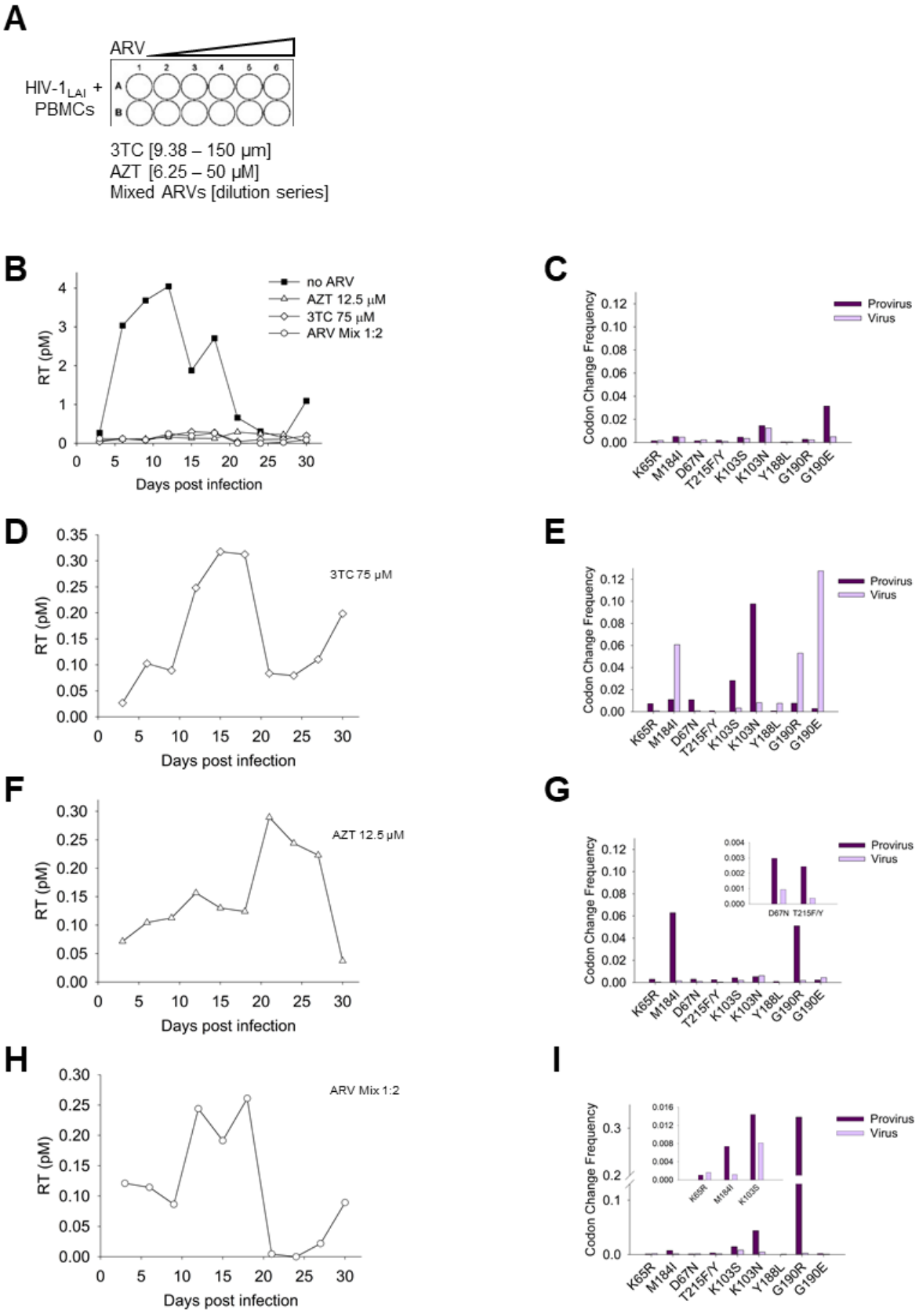

A
ARV
HIV-1LAI + PBMCs
1 2 3 4 5 6
A
B
3TC [9.38 – 150 µm]
AZT [6.25 – 50 µM]
Mixed ARVs [dilution series]
B
no ARV
AZT 12.5 µM
3TC 75 µM
ARV Mix 1:2
RT (pM)
Days post infection
C
Provirus
Virus
Codon Change Frequency
K65R M184I D67N T215F/Y K103S K103N Y188L G190R G190E
D
3TC 75 µM
RT (pM)
Days post infection
E
Provirus
Virus
Codon Change Frequency
K65R M184I D67N T215F/Y K103S K103N Y188L G190R G190E
F
AZT 12.5 µM
RT (pM)
Days post infection
G
Provirus
Virus
Codon Change Frequency
D67N T215F/Y
K65R M184I D67N T215F/Y K103S K103N Y188L G190R G190E
H
ARV Mix 1:2
RT (pM)
Days post infection
I
Provirus
Virus
Codon Change Frequency
K65R M184I K103S
K65R M184I D67N T215F/Y K103S K103N Y188L G190R G190E

Figure 2.3 Drug resistant HIV-1 can be isolated from infected PBMC cultures. (A) PBMCs were cultured for 30 d in with no drug or in the presence of increasing amounts of 3TC (9.38-150 µM), AZT (6.25-50 µM), or Mixed ARVs. (B) Virus replication was measured using an RT assay and showed that the virus replication peaked in the PBMC culture at day 12. Drug resistant variants were also isolated, although the level of virus in the supernatant determined by the RT assay was approximately 10-fold less. (C) Drug resistance mutations were present in both provirus DNA and virus gRNA. (D) Virus resistant to 75 µM 3TC shown in panel A, but on an expanded scale. (E) Virus resistant to 3TC had the corresponding drug resistance mutation at M184I in gRNA, along with drug resistance mutations for other ARVs. (F) Virus resistant to 12.5 µM AZT shown in panel A, but on an expanded scale. (G) Virus resistant to AZT had the corresponding drug resistance mutations at D67N and T215F/Y, along with drug resistance mutations for other ARVs. (H) Virus resistant to a 1:2 dilution of Mixed ARVs (TDF, 3TC, EFV) shown in panel A, but on an expanded scale. (I) Virus resistant to TDF, 3TC, and EFV had corresponding mutations in K65R, M184I, and K103S/N, respectively, along with drug resistance mutations for other ARVs.

2.4.5 Exposure of HIV-1 to A3F and A3G in a single-cycle of replication produces AZT resistant viruses

We next investigated if exposure to A3F and A3G in a single round of replication could introduce drug resistant mutations in otherwise replication competent viral genomes. For this experiment, we used U87 CD4+ CXCR4+ cells and infected them with HIV-1_{LAI}. Two stocks of HIV-1_{LAI} were prepared, one stock was produced in the presence of A3F and A3G so that the A3s were encapsidated into the virions (see Materials and Methods). The other stock was produced in the absence of A3F and A3G. Infection of U87 CD4+ CXCR4+ cells with these virus particles would result in A3 mediated mutagenesis occurring only in the first round of reverse transcription and proviral DNA synthesis, but the cultures were kept for 30 days to determine the fate of the A3 induced mutations on drug resistance to AZT, 3TC, and the Mixed ARV TLE regimen (Figure 2.4A). In the absence of any ARVs, the RT activity of the U87 CD4+ CXCR4+ cells under both No A3 and A3F/A3G conditions had similar profiles and peaked at day 6 (Figure 2.4B). Sequencing of the proviral DNA and cDNA produced from gRNA demonstrated the existence of drug resistant mutations, particularly M184I, K103S/N, and G190R/E (Figure 2.4C-D). The mutation profiles of the No A3 and A3F/A3G condition were very similar, suggesting that the majority of mutations were made from RT-induced error or that RT and A3F/A3G have overlapping mutational hot spots (Figure 2.4C-D).

To discern the impact of RT and A3F/A3G, we conducted 4 independent experiments for U87 CD4+ CXCR4+ cells grown in the presence of 3TC or AZT. The experiments were then scored for whether drug resistance viruses were obtained and if it occurred in the No A3 or A3F/A3G condition (Table 2.2). We found that for 3TC resistance, out of 4 experiments, 3TC resistance arose once in the No A3 condition and once in the A3F/A3G condition (Table 2.2). Although the drug resistant variants that grew from viruses that had packaged A3F/A3G arose early at day 9, they could not maintain growth in culture and by day 12 the RT activity in the supernatant was lost (Figure 2.4E). In addition, the 3TC resistance was at 18.7 μM, less than the 75 μM that arose in the PBMC culture (Figure 2.3D and Figure 2.4E). In contrast, the virus that had no A3 packaged resulted in robust drug resistant viruses emerging from the culture at day 18 in the presence of 75 μM 3TC (Figure 2.4B). Accordingly, the drug resistant viruses from the No A3 condition had very high levels of the M184I mutation in proviral and cDNA isolated from gRNA (Figure 2.4F). In contrast the A3F/A3G exposed viruses had lower amounts of the M184I,

although the mutation was present in cDNA produced from gRNA (Figure 2.4G). The 3TC resistance is also strengthened by the K65R mutation, which can only be induced by RT and was also found in the A3F/A3G condition (Figure 2.4G). Consistent with the D67N mutation being a hot spot for A3-induced mutagenesis in the presence of Vif, this mutation was found at the highest amount in the proviral DNA, but not in the viral cDNA since there was no AZT selection (Figure 2.4G). When the experiments did use AZT, the only conditions that did produce drug resistant virus was the condition with packaged A3F/A3G where we isolated drug resistant virus for two of the four experiments (Table 2.2). Although we recovered virus in one experiment for the No A3 condition in the presence of AZT, the RT activity was 2.5-fold less than the A3F/A3G condition and we were unable to amplify the virus, which was required to isolate gRNA for cDNA synthesis and sequencing (Figure 2.4H). The sequencing data from the A3F/A3G condition in the presence of 12.5 μM AZT showed that the main mutation mediating the resistance was D67N (Figure 2.4I). Although the T215F/Y mutation that results in stronger AZT resistance was found in gRNA, the D67N mutation was 2.4-fold more frequent, consistent with A3F-mediated mutagenesis (Figure 2.2C and Figure 2.4I). Altogether, these data suggest that 3TC resistance from the M184I mutation is more likely to occur from RT error than A3-mediated mutagenesis, but AZT resistance from the D67N mutation is more likely to occur from A3-mediated mutagenesis than RT.

We also examined if exposure to A3F/A3G in a single round of replication could accelerate the development of viruses resistant to the Mixed ARV TLE regimen. We found that the No A3 condition had more robust RT signal at day 15 than the A3F/A3G exposed virus which had 3.5-fold less RT signal and did not become detectable until day 24 (Figure 2.4J). Proviral DNA and gRNA was isolated from each of these drug resistant virus populations and sequenced. For both the No A3 and A3F/A3G conditions, the same drug resistant mutations required for viral replication in the presence of the Mixed ARVs were found in the gRNA, which were K65R, M184I, K103S/N, and G190R/E. Although the codon mutation frequencies were higher in the proviral DNA for the No A3 condition, the virus from the No A3 and A3F/A3G conditions contained similar levels of mutations (Figure 2.4K-L).

Although the A3s could contribute to drug resistant virus variants arising for 3TC and Mixed ARV treatments, the RT signal obtained was less than conditions with No A3 (Figure 2.4E, J). Even with AZT, the RT signal was at least 3-fold less than virus isolated in the absence of A3s

for the other single drug experiment with 3TC (Figure 2.4E, H). These data suggest that A3 enzymes can contribute to drug resistance, but that there may be additional mutations occurring from A3-mediated mutagenesis that compromises the replication capacity of the virus.

2.4.6 Drug resistant viruses derived from A3-mediated mutagenesis have less replication capacity

To examine if there are consequences of A3-mediated mutagenesis on virus replication capacity we tested for emergence of drug resistant variants using the genetically related CEM-SS and CEM cells that have nearly undetectable expression and normal A3F and A3G expression, respectively[230]. We genotyped the A3F of the CEM cells and found that they have the commonly occurring heterozygous A3F genotype of 108S/231I and 108A/231V[160,212]. In treatment naïve cohorts, the A3F 108A/231V genotype has been associated with lower set-point viral load and slower rate of progression to AIDS that appeared to be due to partial resistance to Vif[160]. These CEM-SS/CEM cells were used to enable differentiation between A3- and RT- induced mutations while having continuous encapsidation of endogenously expressed A3s from CEM cells (Figure 2.5A). Without any drug selection, the virus produced from CEM-SS and CEM cells both had peak RT activity after 24 to 27 days (Figure 2.5B). The peak RT activity from the CEM cells was 1.6-fold less than from CEM-SS cells (Figure 2.5B). Sequencing of proviral DNA and cDNA isolated from gRNA showed that although there were more drug resistance mutations in the CEM-SS conditions, the types of mutations between the no A3 (CEM-SS) and A3 (CEM) conditions were similar types, suggesting they resulted from RT error (Figure 2.5C-D).

To examine specifically the function of A3s in emergence of drug resistance, we grew CEM-SS and CEM cells in either 3TC or AZT. Data from infections in U87 CD4+ CXCR4+ cells suggested that the M184I mutation in replicating viruses was induced more efficiently by RT error than A3-mediated mutagenesis (Figure 2.4). In support of this line of reasoning, we isolated viruses that had high RT activity from CEM-SS cells in the presence of 150 μM 3TC (Figure 2.5E). Although we could isolate viruses from CEM cells at 150 μM 3TC, the level of RT activity was 24-fold lower than viruses from CEM-SS cells (Figure 2.5E). The M184I was present in gRNA for both CEM-SS and CEM cells, but the frequency was 2.5-fold higher in virus from CEM-SS cells (Figure 2.5F-G). There were also multiple other drug resistance mutations present, such as K103S/N in CEM-SS cells, consistent with RT-induced mutations, and D67N in CEM cells,

consistent with A3-mediated mutagenesis (Figure 2.2C and Figure 2.5F-G). Interestingly, the G190R occurred at a high frequency in proviral DNA of viruses from CEM-SS cells, but not CEM cells, even though this codon has an A3 deamination motif and was a preferred deamination site for A3G in single-cycle infectivity assays (Figure 2.2F and Figure 2.5F-G).

For selection of drug resistant variants in the presence of AZT, consistent with the results from U87 CD4+ CXCR4+ cells, the CEM cells resulted in AZT resistance at 25 μM AZT, but the CEM-SS cells resulted in resistance only at 6.25 μM AZT (Figure 2.4 H-I and Figure 2.5H, J). We also obtained virus from CEM cells resistant to 6.25 μM AZT, but the RT activity had ~40-fold less RT activity than drug resistant virus from CEM-SS cells (Figure 2.5H). Even the CEM derived drug resistant virus isolated at 25 μM AZT had 15-fold less RT activity than CEM-SS derived virus resistant to 6.25 μM AZT (Figure 2.5H, J). All the AZT resistance was from the D67N mutation in the gRNA, since we did not recover any T215F/Y resistance mutations in gRNA, although it was found in the proviral DNA of viruses isolated from both CEM-SS and CEM cells (Figure 2.5 I, K). These data collectively show that although A3s can induce drug resistance, the resulting viruses had less replication capacity than drug resistant viruses derived from RT-induced mutations alone.

Table 2.2 Emergence of 3TC or AZT resistant viruses in the absence or presence of encapsidated A3F/A3G.

3TC	Drug resistant virus	
	No A3	Packaged A3F/G
Experiment 1		Yes
Experiment 2		
Experiment 3		
Experiment 4	Yes	

AZT	Drug resistant virus	
	No A3	Packaged A3F/G
Experiment 1		
Experiment 2		Yes
Experiment 3		
Experiment 4		Yes

Experiment was carried out as detailed in Figure 2.4A sketch for 3TC and AZT. Occurrence of drug resistant viruses was scored for four independent experiments. No notation means that no drug resistant virus was isolated.

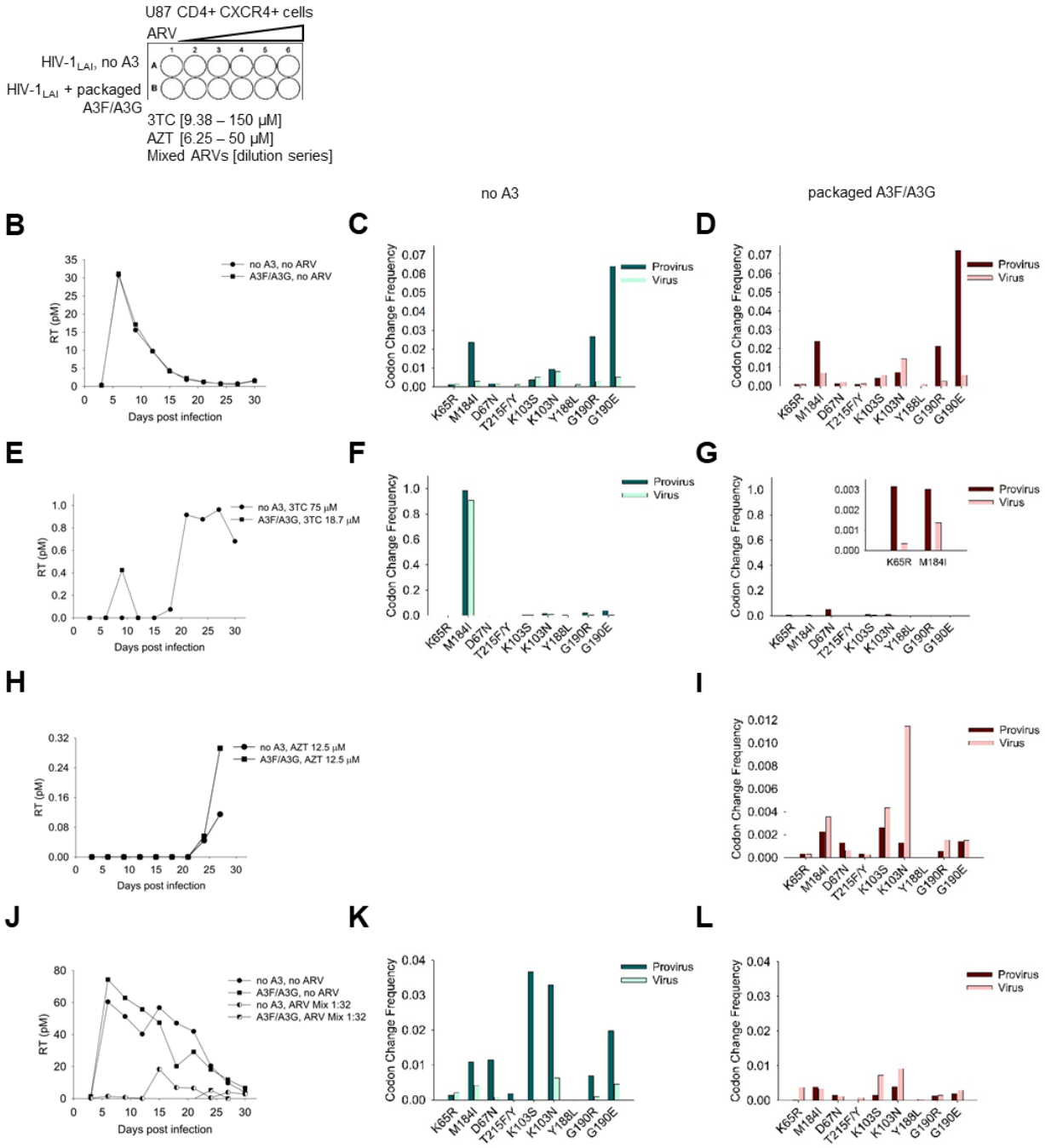

U87 CD4+ CXCR4+ cells
ARV
HIV-1LAI, no A3
HIV-1LAI + packaged A3F/A3G
3TC [9.38 – 150 µM]
AZT [6.25 – 50 µM]
Mixed ARVs [dilution series]
no A3
packaged A3F/A3G
B
C
D
E
F
G
H
I
J
K
L
RT (pM)
Days post infection
no A3, no ARV
A3F/A3G, no ARV
no A3, 3TC 75 µM
A3F/A3G, 3TC 18.7 µM
no A3, AZT 12.5 µM
A3F/A3G, AZT 12.5 µM
no A3, ARV Mix 1:32
A3F/A3G, ARV Mix 1:32
Codon Change Frequency
Provirus
Virus
K65R
M184I
D67N
T215F/Y
K103S
K103N
Y188L
G190R
G190E

Figure 2.4 Exposure of WT HIV-1 to A3G and A3F for a single round of replication can generate drug resistant viruses. (A) U87 CD4+ CXCR4+ cells were cultured for 30 d in with no drug or in the presence of increasing amounts of 3TC (9.38-150 μM), AZT (6.25-50 μM), or Mixed ARVs. (B) Virus replication was measured using an RT assay and showed that the virus replication peaked in the U87 CD4+ CXCR4+ culture at day 6. (C-D) Drug resistance mutations were present in both provirus DNA and virus gRNA for the No A3 and coexpressed A3F and A3G (A3F/A3G) condition. (E) Virus resistant to 75 μM 3TC (No A3) and 18.7 μM 3TC (A3F/A3G) was isolated. (F-G) Virus resistant to 3TC had the corresponding drug resistance mutation at M184I in gRNA, along with drug resistance mutations for other ARVs. (G) The K65R mutation that provides additional resistance to 3TC was found in the A3F/A3G condition. (H) Virus resistant to 12.5 μM AZT was isolated from the A3F/A3G condition. (I) Virus resistant to AZT had the corresponding drug resistance mutations at D67N and T215F/Y, along with drug resistance mutations for other ARVs. (A-I) Data was obtained from the experiments listed in Table 2.2. Virus replication in the absence of drug was determined for each experiment in the No A3 and A3F/A3G condition, but only one representative curve is shown for each condition in panel B. (J) Virus resistant to a 1:32 dilution of Mixed ARVs (TDF, 3TC, EFV) was isolated, but similar to the other drug resistant viruses, the amount of RT activity in the supernatant was less than without the Mixed ARV. (K-L) Virus resistant to TDF, 3TC, and EFV had corresponding mutations in K65R, M184I, and K103S/N, respectively, along with drug resistance mutations for other ARVs.

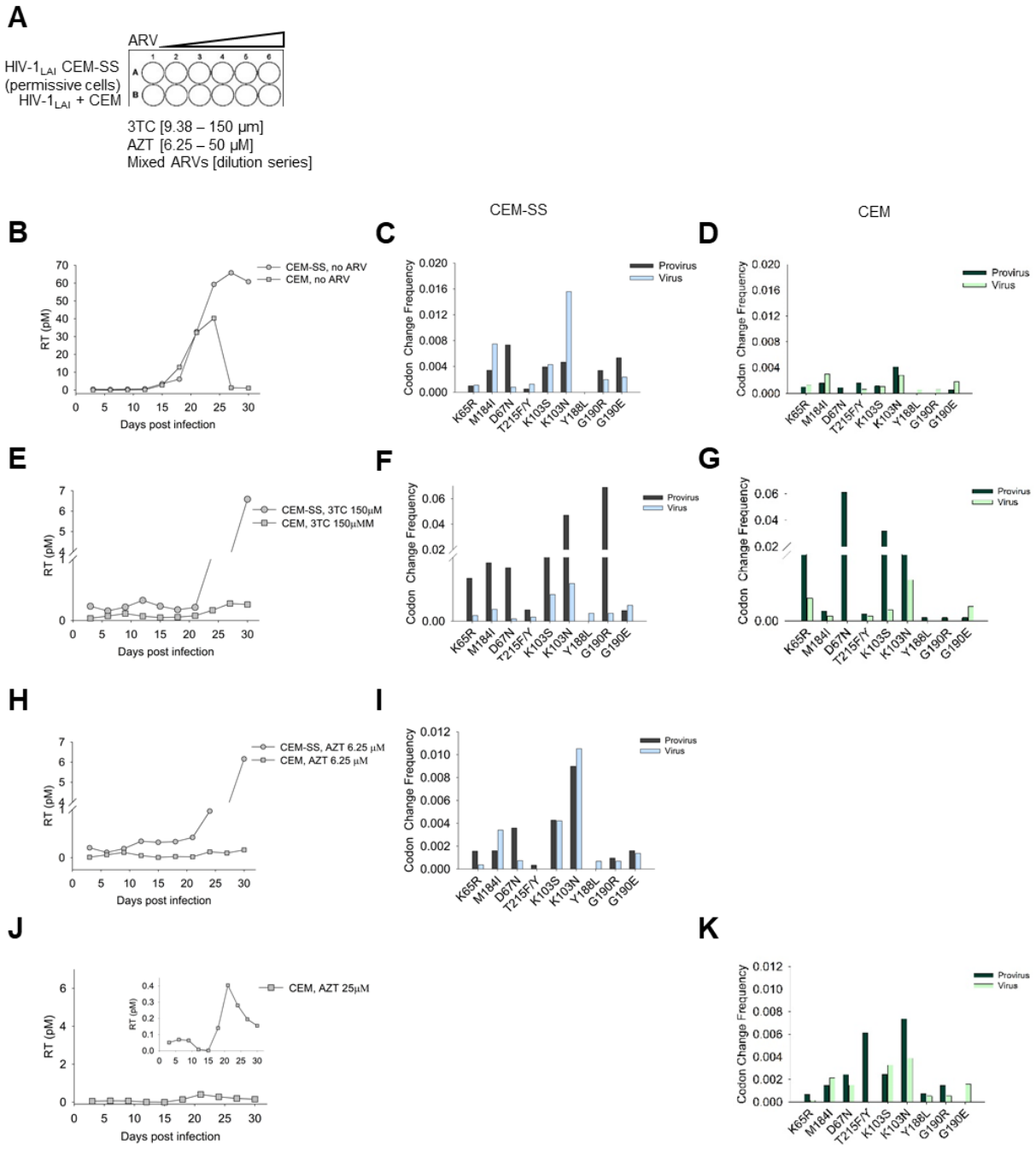

A
ARV
HIV-1LAI CEM-SS
(permissive cells)
HIV-1LAI + CEM
1 2 3 4 5 6
A
B
3TC [9.38 – 150 µm]
AZT [6.25 – 50 µM]
Mixed ARVs [dilution series]
CEM-SS
CEM
B
C
D
E
F
G
H
I
J
K
RT (pM)
Days post infection
CEM-SS, no ARV
CEM, no ARV
CEM-SS, 3TC 150µM
CEM, 3TC 150µMM
CEM-SS, AZT 6.25 µM
CEM, AZT 6.25 µM
CEM, AZT 25µM
Codon Change Frequency
Provirus
Virus
K65R
M184I
D67N
T215F/Y
K103S
K103N
Y188L
G190R
G190E

Figure 2. 5 Drug resistant viruses induced by A3G- or A3F- mediated mutations have less replication capacity than drug resistant viruses that arise from RT-induced error. (A) CEM-SS (permissive) or CEM cells (non-permissive) cells were cultured for 30 d in with no drug or in the presence of increasing amounts of 3TC (9.38-150 μM) or AZT (6.25-50 μM). (B) Virus replication was measured using an RT assay and showed that the virus replication peaked in the CEM-SS culture at day 27 and the CEM culture at day 24. (C-D) Drug resistance mutations were present in both provirus DNA and virus gRNA for the No A3 and coexpressed A3F and A3G (A3F/A3G) condition. (E) Virus resistant to 150 μM 3TC was isolated fro mboth CEM-SS and CEM cultures (F-G) Virus resistant to 3TC had the corresponding drug resistance mutation at M184I in gRNA, along with drug resistance mutations for other ARVs for (F) CEM-SS and (G) CEM cells. (H) Virus resistant to 6.25 μM AZT was isolated from cultures of CEM-SS and CEM cells. (I) Virus resistant to AZT from the CEM-SS culture was sequenced and had the corresponding drug resistance mutations at D67N and T215F/Y, along with drug resistance mutations for other ARVs. (J) Virus resistant to 25 μM AZT was isolated from CEM cell culture. (K) Virus resistant to AZT from the CEM culture was sequenced and had the corresponding drug resistance mutations at D67N and T215F/Y, along with drug resistance mutations for other ARVs.

2.5 Discussion

The extent to which A3-mediated mutagenesis can contribute to the development of drug resistance is not clear. This could be due to different results from several studies using different experimental systems. In particular, some key limitations from previous studies were that only A3G instead of multiple A3s was assessed, the bias of RT to induce G→A mutations was not considered, only one antiviral drug was used, and only viral recombination with ΔVif HIV or HIV with partially inactive Vif was considered. To overcome these past limitations, we used WT HIV and assessed the contribution of A3F and A3G on drug resistance to multiple drugs, alone and in combination, in comparison to conditions with no A3 enzymes. We determined that all the G→A RT-induced mutations were in a G→A context that matched that of A3F or A3G, confounding the interpretation of many previous studies. Of utmost importance, we found that expression of both A3F and A3G in the HIV producer cells resulted in a protection of A3F from complete Vif-mediated degradation and resulted in a proviral mutational load from approximately equal contributions of A3F and A3G, despite A3F previously being found less active in a ΔVif HIV infection[133,217,218,233,234]. Altogether, in separate cell lines we found the consistent result that although A3 enzymes could induce drug resistant mutations that resulted in viral replication under drug selection pressure, those viruses had less replication capacity than drug resistant viruses induced by RT alone. Thus, the contribution of A3-mediated mutations to the emergence of drug resistance is unlikely to be significant. These data support a model in which A3 enzymes cooperate to ensure viral suppression rather than viral evolution.

Consistent with previous studies, we also found in our mutation analysis that A3G could induce the M184I mutation in proviral DNA[150,216]. We found in a single-cycle infectivity assay with WT HIV that A3G could induce this mutation in the absence of 3TC at a higher level than A3F and RT alone (Figure 2.2). However, despite this ability we found that A3G was less able to produce drug resistant virus that could replicate as robustly as drug resistant viruses that emerged in the absence of A3G and A3F (Figure 2.4 and Figure 2.5). Further, in the U87 CD4+ CXCR4+ cells, the virus isolated from day 6 had approximately equal levels of the M184I mutation between the No A3 and A3F/A3G condition, in contrast to the single-cycle infectivity assay, and suggesting that RT over several replication cycles can also induce the M184I mutation (Figure 2.2 and Figure 2.3). Nonetheless, that we did isolate replication competent 3TC resistant virus in the presence of

A3F and A3G underscores that viral inactivation after A3-mediated mutagenesis during a WT HIV infection is not guaranteed (Figure 2.4 and Figure 2.5). There are several missense mutations that can be induced by A3 enzymes that do not inactivate key enzymes such as protease[77,133,134]. However, we observed that sublethally mutated and drug resistant proviral genomes are unlikely to replicate as well as WT HIV. As a result, the differentiation between sublethal mutagenesis and viral suppression appears to be the number of mutations, in addition to the nature of the mutations. A3G has been shown to be very efficient at inactivation of ΔVif HIV with some reports suggesting that even a single encapsidated A3G molecule can induce lethal mutations that inactivate the virus[131,133,217,218,233-236]. Even though A3G is more efficient at inducing mutations than A3F in a ΔVif HIV infection, the A3G is also more sensitive to Vif (Figure 2.1). In our study, A3G was always co-expressed with A3F. This may have resulted in additional mutations that led to drug resistant viruses with less replication capacity[126,144,237-239].

Although A3F has been implicated more than A3G for inducing genetic diversity, no studies have directly tested the ability of A3F to promote HIV drug resistance[133,217,219]. Here we showed that A3F can preferentially induce the D67N mutation that results in AZT resistance and this occurred in the presence of Vif (Figure 2.3, Figure 2.4, Figure 2.5). In U87 CD4+ CXCR4+ cells with virus exposed to A3F/A3G for only one round of replication, the A3F/A3G condition resulted in AZT resistant virus with more replication capacity than the No A3 condition for two of the four experimental replicates (Table 2.2, Figure 2.3). Despite this induced resistance, we also found that in CEM cells where there is continual encapsidation of A3F and A3G, the resulting viruses still became resistant to AZT, but had less replication capacity than the AZT resistant viruses that emerged from CEM-SS cells in the absence of any A3s (Figure 2.5). Our data demonstrated that the coordinate expression of A3G with A3F results in more encapsidation of A3F due to protection from Vif-mediated degradation, resulting in almost equal contribution of mutations from A3F and A3G in a WT HIV infection (Figure 2.1, Figure 2.2). This suggests that the combination of GG→AG and GA→AA mutations could have a higher inactivation potential than each alone in a WT HIV infection[236,237]. This may be due to differences in enzyme processivity that result in A3F inducing more mutations that are widely spaced, potentially at least partially inactivating more target proteins, and A3G inducing clusters of mutations that have high inactivation potential if they fall within conserved regions or enzyme active sites[133].

Our study also addressed the contribution of RT-induced G→A mutations and whether the number of A3-induced mutations could exceed the RT error rate in a single round of replication. Several studies have assessed levels of A3-induced mutagenesis using proviruses recovered after unknown numbers of multiple rounds of virus replication and sometimes from infections with naturally occurring defective Vif viruses[119,150,152,172,239]. If the proviral DNA is "hypermutated", then recombination with a WT or sublethally mutated virus would be required to recover drug resistant mutations. While this may efficiently occur in cell culture, it appears less likely in a clinical infection, suggesting that this is not a main route of A3-induced drug resistance, if it occurs[150,152]. Rather, the sublethal mutagenesis that would occur in the first round of replication would be the most likely mechanism for A3 enzymes to contribute to drug resistance. In this situation, we found evidence of potential sublethal mutagenesis since A3-induced mutations decreased from ~20-fold above RT error to only 4-fold above RT error (Figure 2.2). These mutations did not cause any decrease in viral replication in the absence of drug selection (Figure 2.1). However, in the presence of drug selection where only select viruses could replicate, the effect of the A3-induced mutations was visible because the virus replication, as measured by an RT assay, consistently resulted in less signal in the presence of AZT, 3TC, and the combined ARVs (Figure 2.4 and Figure 2.5). Importantly, in a spreading infection we found that the chance for RT to induce the M184I mutation in comparison to viruses exposed to A3G and A3F for a single round of replication was equal, but under drug selection pressure, the no A3 condition produced a much stronger pool of 3TC resistant virus (Figure 2.4 and Figure 2.5). These data suggest that previous studies have underestimated the ability of RT alone to induce this mutation. Particularly since all of our quantified RT induced G→A mutations were within a relevant A3 mutation context (Figure 2.2), consistent with analysis of the mutational bias of HIV-1 in proviral DNA from HIV+ individuals[226] and studies characterizing RT fidelity[90,148,229]. Only the resistance to AZT was consistently induced by A3 enzymes, likely A3F, but AZT resistant viruses continually exposed to A3G and A3F had lower replication capacity than viruses not exposed to A3G and A3F (Figure 2.5). With a combined ARV treatment, the single round of exposure to A3s enabled resistance, but it was less robust than RT alone (Figure 2.4). Altogether, the data suggest that the impact of A3F and A3G on drug resistance is less than RT alone.

Previously published studies found different results than ours and suggest that A3G alone can induce the M184I mutation required for virus resistance to 3TC. In Mulder *et al.*, a previously

identified HIV-1 with point mutations in Vif that caused less A3G degradation was used to generate a pool of proviruses[150]. These proviruses, similar to the ones isolated in our study, contained the M184I mutation prior to any drug exposure (Figure 2.2)[150]. Also similar to our results in using ΔVif HIV, the A3G was able to induce the M184I mutation to a much greater extent than A3F and RT (Figure 2.2)[150]. However, in the Mulder et al. study, the large majority of provirus was inactive and in order to select for resistant viruses, the provirus had to be co-infected with WT HIV[150]. With no WT virus added to the cells for recombination as in the Mulder *et al*. study, we found that A3G was not inducing 3TC drug resistance (Figure 2.4)[150]. In the U87 CD4+ CXCR4+ cells, we found that although the frequency of the M184I mutation was similar between the No A3 and A3F/A3G conditions in the absence of 3TC, in the presence of 3TC the No A3 condition had 300- and 600- fold higher frequencies of the M184I mutation in provirus and virus, respectively (Figure 2.4). While the scenario proposed by Mulder *et al*., where individuals have viruses with partially defective Vifs does occur, the extent to which superinfection with WT HIV and recombination with defective viruses occurs appears to decrease the possibility of A3G-induced drug resistance arising by this mechanism[120,152,159]. In contrast, a study by Kim *et al*. used an experimental system similar to ours with WT HIV and CEM-SS cells, but obtained results indicating that only CEM-SS cells that stably expressed A3G, but not CEM-SS cells alone, could induce 3TC drug resistance[216]. Although we did recover 3TC resistant virus from CEM cells, it replication ability was 24-fold lower than CEM-SS cells (Figure 2.5). The key difference between our studies is that we also had A3F present, which may result in higher levels of mutagenesis and ensure viral inactivation. In contrast, Kim *et al*. found no effect of A3G on viral fitness, in agreement with an early work from Sadler *et al*.[216,240] While this is in agreement with our data showing equal virus replication in the presence and absence of A3G and A3F without drug selection, it appears that the drug resistant viruses that become selected from the viral population have additionally acquired collateral mutations that impair virus replication (Figure 2.4 and Figure 2.5).

Despite different results from some cell-based experiments, our data are in agreement with several clinical and *in silico* studies regarding the effects of collateral A3-induced mutations. An early study by Jern *et al*. found similar results through *in silico* simulation analysis that concluded A3G deaminations have been a factor in HIV evolution, but only a small proportion of the deaminations could be responsible for drug resistance mutations, indicating that there is not always

a direct connection between resistance and A3-mediated mutagenesis[172]. This was supported by a study by Delvis-Frankenberry *et al*. that suggested hypermutated proviruses occur, but are rarely rescued by recombination and that sublethal mutagenesis for A3G and A3F is negligible since it occurs at a rate less than the RT mutations[152]. While our single-cycle infectivity assays found a larger number of mutations due to A3G and A3F in the presence of Vif (Figure 2.2), the overall conclusions were the same; that the effect of A3-mediated mutagenesis appeared to be either be very low and insignificant or high enough to inactivate or partially suppress viral replication[152]. In particular, several studies have found stop codons almost concomitant with A3-mediated mutagenesis, even with WT HIV infections and in clinical studies[152,224,227,228]. In particular, the M184I mutation causes RT to be less processive, although this can be compensated for with an E138K mutation[241-246]. Nonetheless, the M184I mutation slows the synthesis of DNA and enables more time for A3-mediated mutagenesis, resulting in more mutations and making viral inactivation more likely and the contribution of A3 enzymes to drug resistance less likely[225]. The level of A3-induced mutations differs for various anatomical compartments, but consistently A3 mutated proviral genomes also contain a high proportion of inactivating mutations[159]. Further, during a WT HIV infection, even the E138K mutation is more likely to be induced by RT, due to its G→A error bias[226], consistent with our sequencing data (Figure 2.2). Altogether, the data suggest that there is a higher likelihood for RT to induce drug resistant viruses with high replication capacity, rather than A3 enzymes. Collectively the data from our study and others suggests that there is a large dependence on the environmental selection pressures for whether A3-induced mutations result in replication competent drug resistant virus. This also appears to be the case for the involvement of A3 enzymes in the CTL response. Where some studies have found that A3-induced mutations can increase CTL recognition of HIV infected cells others have found that A3-induced mutations predominantly decrease the recognition of HIV infected cells[173,174,214,215,220-222]. Either way, it has also been found that A3-induced mutations usually coincide with functional viral inactivation, but not disruption of transcription[173,214]. Altogether, the effect of A3 enzymes can bias the immune response to target cells carrying defective proviruses, rather than causing direct immune escape *per se*[214]. This broader view suggests that what is more important than whether the CTL response is evaded or accentuated is whether the provirus is actively able to replicate or not. When this broader view is applied to A3-induced drug resistance, the data from our work and others suggest that although A3 enzymes can induce drug resistance in proviral DNA, the chance of these viruses

being replication competent and having a competitively strong replication ability is less likely than mutations induced by RT alone.

2.6 Acknowledgements

We acknowledge Dr. Janet Hill and Champika Fernando (Western College of Veterinary Medicine, University of Saskatchewan) for assistance with Illumina sequencing.

Table 2.3 List of primers

Primers used for amplifying cDNA and the specific regions on *pol* harbouring the potential drug resistance mutations	
Ext F p7.9	AAAGCATTGGGACCAGGAGCGACACTAGAAGARATGATGACAGCATGYCA
Half RT	TATTTCTGCTATTAAGTCTTTTGATGGGTCA
Ext F p7.10	GGAGCGACACTAGAAGAAATGATGACAGCATGYCARGGAGTRGGRGGRCY
Half Pol	TCTGCCAGTTCTAGCTCTGCTT
D67N_F	CCAGTAAAATTAAAGCCAGGAATG
D67N_R	AGTATACTTCCTGAAGTCTTCATCT
E138K_F	AATAAGAGAACTCAAGACTTCTGG
E138K_R	TGTTCTATGCTGCCCTATTTCTA
M184I_F	CATACCTAGTATAAACAATGAGACAC
M184I_R	CTGTCTTTTTCTGGCAGCACTAT
D67N_F_MS	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCAGTAAAATTAAAGCCAGGAATG
D67N_R_MS	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGTATACTTCCTGAAGTCTTCATCT
E138K_F_MS	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAATAAGAGAACTCAAGACTTCTGG
E138K_R_MS	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGTTCTATGCTGCCCTATTTCTA
M184I_F_MS	TCGTCGGCAGAGTCAGATGTGTATAAGAGACAGCATACCTAGTATAAACAATGAGACAC
M184I_R_MS	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTGTCTTTTTCTGGCAGCACTAT

3.0. POLYMORPHISMS OF THE CYTIDINE DEMINASE APOBEC3F COOPERATE TO RESTRICT HIV-1

Nazanin Mohammadzadeh[a], Tyson B. Follack[a], Robin P. Love[a], Kris Stewart[b,c,d], Stephen Sanche[b, d], and Linda Chelico[a, d]

[a] University of Saskatchewan, Biochemistry, Microbiology, and Immunology, College of Medicine, Saskatoon, Saskatchewan, Canada

[b] University of Saskatchewan, Department of Medicine, College of Medicine, Saskatoon, Saskatchewan Canada

[c] Saskatchewan Infectious Disease Care Network, Saskatoon, Saskatchewan, Canada

[d] Saskatchewan HIV/AIDS Research Endeavour, Saskatoon, Saskatchewan, Canada

This chapter was previously published as Mohammadzadeh et al., Virology, 527: 21-31 (2018).

https://www.sciencedirect.com/science/article/pii/S0042682218303453

Key Words: HIV-1; APOBEC3; Vif; restriction factors; infectivity; deamination; mutagenesis

All experiments and data analysis in this chapter were performed by N.M., R.P.L. T.B.F., and L.C. Figures 3.1 and 3.2 were performed by N.M. Figure 3.3 was performed by T.B.F. Figure 3.4 was performed by R.P.L. T.B.F., and L.C. N.M., T.B.F., and L.C. conceived and designed the experiments.

3.1 Abstract

The APOBEC3 enzyme family are host restriction factors that induce mutagenesis of HIV proviral genomes through the deamination of cytosine to form uracil in nascent single-stranded (-)DNA. HIV suppresses APOBEC3 activity through the HIV protein Vif, that induces APOBEC3 degradation. Here we compared two common polymorphisms of APOBEC3F. We found that although both polymorphisms have HIV-1 restriction activity, APOBEC3F 108A/231V can restrict HIV-1 ΔVif up to 4-fold more than APOBEC3F 108S/231I and is partially protected from Vif-mediated degradation. This resulted from higher levels of steady state expression of APOBEC3F 108A/231V. Individuals are commonly heterozygous for the APOBEC3F polymorphisms and these polymorphisms formed in cells, independent of RNA, hetero-oligomers between each other and with APOBEC3G. Hetero-oligomerization with APOBEC3F 108A/231V resulted in partial stabilization of APOBEC3F 108S/231I and APOBEC3G in the presence of Vif. These data demonstrate functional outcomes of APOBEC3 polymorphisms and hetero-oligomerization that affect HIV restriction.

3.2 Importance

APOBEC3 enzymes are a barrier to HIV infection and suppress HIV replication by inducing mutagenesis of HIV proviral DNA. To counteract this host restriction system HIV uses Vif to induce degradation of relevant APOBEC3 enzymes. Although there is a constant battle between host and virus for suppression of replication and a counteraction, the HIV mutation rate far exceeds that of humans and facilitates a faster rate of adaptation. As a result, protective measures such as having multiple host genotypes can guard populations by slowing the rate of viral adaptation. We examined if two common genotypes of APOBEC3F were an example of multiple genotypes that can help protect the host. We found that the two common APOBEC3F genotypes most often occur as heterozygous alleles and that the one allele that was more protective could stabilize APOBEC3F and APOBEC3G in cells to increase virus encapsidation and partially protect against Vif mediated degradation.

3.3 Introduction

APOBEC3 (A3) enzymes are a family of deoxycytidine deaminases that can act as host restriction factors for HIV [17,92]. The A3 enzymes deaminate cytosine in single-stranded (ss) DNA

to form uracil [25,247-249]. Of the seven human A3 enzymes, there are five that can restrict the replication of HIV in CD4+ T cells through the formation of uracils in (-) strand proviral DNA. These enzymes, A3D, A3F, A3G, A3H (haplotypes II, V, and VII), and A3C (S188I) restrict HIV with varying efficiencies [23,110,111,126,131,209,250,251]. In an HIV infected CD4+ T cell, these enzymes can become encapsidated into assembling virions through an interaction with cellular or HIV genomic RNA that is also bound to HIV Gag [117,183]. After virion maturation and infection of a target cell, the encapsidated A3 enzymes can access single-stranded HIV (-) DNA when it is synthesized and exposed through reverse transcription and RNase H activity [125]. The cytosines that become deaminated to uracil are copied by the reverse transcriptase during (+) DNA synthesis and result in guanine to adenine (G→A) mutations on the coding strand [25,125,247-249]. These mutations can induce DNA repair enzyme-mediated degradation of proviral DNA or result in functional inactivation of integrated proviral DNA [205]. HIV-1 Vif suppresses the encapsidation and activity of A3 enzymes by mediating their polyubiquitination and proteasomal degradation [26,29-32,252,253]. HIV-1 Vif hijacks a Cullin 5 E3 ubiquitin ligase complex and acts as the substrate receptor by replacing the host protein SOCS2 [29]. This enables HIV-1 Vif to interact directly with components of the ubiquitin ligase complex, Elongin C and Cullin 5 [33]. HIV-1 Vif can also directly interact with A3D, A3F, A3G, A3H, and A3C, which induces their ubiquitination and proteasomal degradation [185,254-256]. Obligatory for stability, Vif also interacts with the transcription cofactor CBF-β [27,257].

Despite the presence of Vif, A3 enzymes are still able to be encapsidated into HIV virions and induce mutagenesis [156,195-204,258]. Although translation of Vif mRNA results in high levels of Vif in cells before virus particle assembly, which depletes A3s from infected CD4+ T cells, the induced degradation of A3s is not complete [259]. Evidence of this comes from high levels of A3 induced stop codons and missense mutations in HIV proviral DNA in acute and chronic proviral DNA isolated from HIV+ individuals [199]. Sequencing of integrated proviral genomes from multiple studies have revealed that the mutations are present on the (+) DNA in both a 5′GG→5′AG context and a 5′GA→5′AA context [127,156,195-204,258].Mutations in the 5′GG context originates primarily from A3G that deaminates preferentially in 5′C<u>C</u>A motifs on the (-) DNA (underlined C deaminated) [125]. This sequence context most often results in mutations that introduce stop codons and mutations at glycines [77]. Mutations in the 5′GA context originates primarily from A3D, A3F, A3H, and A3C (S188I) due to a preference to deaminate in 5′T<u>C</u> motifs on the (-) DNA

[126,127,137,260]. Mutations at the 5′TC motif can also introduce stop codons and missense mutations in a wider variety of amino acids [77]. The fate of these mutations may inactivate HIV-1 or may lead to evolution in various forms, such as drug resistance, CTL escape, and changes in co-receptor usage [150,171,174,215,219,261]. However, depending on the experimental system some studies have found that A3 enzymes do not contribute to drug resistance and may enhance CTL recognition [152,172,221,222,261]. Although A3G is the most active deaminase in a ΔVif HIV-1 infection, it is also very sensitive to Vif [133,218,262]. Thus, despite there being lower activity from A3D and A3F in ΔVif HIV-1 infections, there is some evidence that A3F is less sensitive to Vif, and as a result, may be a greater contributor to mutagenesis during HIV infections *in vivo* [131,133,136,160,217]. There is also evidence that in chronically infected individuals that HIV-1 Vif can acquire mutations that result in partial activity, thereby enabling A3 encapsidation and mutagenesis [119]. However, it is unlikely that these mutations in Vif are acquired early in infection and since A3-mediated mutagenesis is found in acute infections, the data suggest that other mechanisms can enable A3 encapsidation in the presence of Vif [199].

Resistance to Vif mediated degradation has been characterized for some A3 enzymes. A3H exists as seven major haplotypes in humans with different protein stability, HIV restriction ability, and sensitivity to Vif [111,113,263]. These haplotypes result from a combination of five different single-nucleotide polymorphisms (SNPs) [111,113,263]. The diversity of A3H in the human population necessitates constant adaptation of HIV-1 Vif to A3H, especially since stable and restrictive A3H haplotypes (II, V, and VII) show population stratifications [111,113,127]. These dynamics result in a situation where A3H can act as a transmission barrier for some HIV-1 strains that do not have a Vif capable of inducing A3H degradation [264]. At the very least, this A3H resistance to some HIV-1 Vifs can increase the time until CD4+ T cell counts decrease or until the development of AIDS, while Vif acquires the adaptive mutations to overcome A3H [127]. The number of SNPs in other A3s is significantly lower and most commonly they have been found to result in inactivation of the A3 or no effect [210-212]. Exceptions are A3D and A3C in which SNPs have resulted in increased activity [110,209,212]. Interestingly, despite cell based assays not uncovering differences in a common variant of A3F A108S/V231I, a population based analysis of multiple pre-treatment cohorts of HIV+ individuals showed that the A3F 231V variant was associated with lower viral load and slower rate of progression to AIDS [160,212,233]. The authors suggested that the 231V polymorphism might increase resistance to Vif [160]. Despite this amino acid not being within the canonical Vif binding

interface on A3F, this observation was also previously reported for A3F 108A/231V [126,186,190,265,266].

Here we assessed the A3F 108S/231I and A3F 108A/231V variants for their ability to restrict HIV-1 replication in the presence and absence of Vif. Based on past evidence that A3F can interact with itself and A3G, in an RNA-independent manner [144], we assessed how each A3F variant restricted HIV-1 replication when expressed together and in the presence of A3G. Notably, the SNPs responsible for the A3F 108A/231V and 108S/231I (rs202390, rs2076101) have been identified to commonly occur together in genetic studies [212,233]. Further, analysis of multiple human populations from the Ensembl genome browser (release 94) showed that these SNPs are in high linkage disequilibrium ($r^2 > 0.95$) for European, South Asian, and East Asian populations and linkage disequilibrium also occurs, but at lower frequencies, for Mixed American ($r^2 > 0.87$) and African ($r^2 > 0.35$) populations [267]. As a result, we refer to these variants as A3F 231V or A3F 231I for brevity and because this specific variation was correlated epidemiologically in affecting HIV-1 progression [160,212,233]. Our analysis of these A3F variants found that the A3F 231V is a more efficient restriction factor for HIV-1 than A3F 231I. The increased restriction was due to an increase in A3F steady state levels in cells. This increased steady state level of A3F 231V resulted in a partial protection from Vif-mediated degradation. Notably, population analysis demonstrated that the A3F 231V and A3F 231I alleles occurred most often together, rather than the homozygous alleles. This appeared to have a functional benefit since the interaction of A3F 231V could also increase the Vif resistance of A3F 231I and A3G, providing a possible reasoning for how high numbers of A3 induced mutations may be acquired in HIV+ individuals.

3.4 Results

3.4.1 A3F 231V more efficiently restricts HIV-1 infection than A3F 231I

The A3F 231V allele has been shown to be advantageous in HIV+ individuals, but the reason for this was not fully explored [160]. Here we conducted an assessment of the HIV-1 restriction ability of A3F 231V and A3F 231I. The effect of untagged A3F 231V and A3F 231I on VSV-G pseudotyped ΔVif HIV-1 infectivity was assessed using transient transfection of different amounts of A3F vector in single-cycle replication assays. Over a range of A3F plasmid transfection amounts, the A3F 231V restricted ΔVif HIV-1 1.4-fold (25 ng, $p \leq 0.05$), 1.6-fold (50 ng, $p \leq 0.01$),

and 4.0-fold (100 ng, $p \leq 0.001$) more than A3F 231I (Figure 3.1A). Overall the ability for A3F to restrict ΔVif HIV-1 was still 1.5-fold (25 ng, $p \leq 0.01$), 1.8-fold (50 ng, $p \leq 0.01$), and 13.6-fold (100 ng, $p \leq 0.001$) less than A3G, for the more active A3F 231V variant, consistent with previous reports (Figure 3.1A) [131]. Also consistent with A3F 231V being protective in HIV+ individuals [160], the A3F 231V enabled modest restriction of virus replication in the presence of Vif, which was not observed for A3F 231I or A3G (Figure 3.1B, $p \leq 0.01$ for 50 ng and 100 ng).

To confirm that decreases in infectivity in the presence and absence of Vif correlated with virion encapsidation of the A3 enzymes, the cell lysates and virions from the replication assays were analyzed after proteins were resolved by SDS-PAGE and transferred to membranes for immunoblotting. The native antibody used to detect A3F was only able to detect cellular expression (Figure 3.1C-E) and not virion encapsidation, except at the 100 ng transfection amount (Figure 3.1E and data not shown). These blots showed that the A3F 231V had a higher steady state expression level than A3F 231I in the absence of Vif (5- to 10- fold) and presence of Vif (2.5- to 15- fold) (Figure 3.1C-E), consistent with higher decreases in HIV-1 infectivity in both of these conditions (Figure 3.1A-B). At the highest plasmid transfection amount (100 ng), the A3F variants could be detected in the virus particles. The A3F 231V was encapsidated 5-fold more than A3F 231I in the absence of Vif and in the presence of Vif encapsidation of the A3F 231V, but not the A3F 231I was detected (Figure 3.1E). Collectively, these data suggest that the A3F 231V is partially protected from Vif-mediated degradation. Similar limitations with a native antibody for A3G were identified in that the cellular expression levels were detected with the native antibody, but the encapsidation was not detected until the plasmid transfection amount reached 100 ng, despite evidence of HIV-1 restriction in the single-cycle replication assays (Figure 3.1A and F-H). Although significant amounts of A3G were detected in the virus in the presence of Vif at the higher transfection amount (Figure 3.1H), there was no observable decrease in infectivity (Figure 3.1B), presumably due to Vif-mediated inhibition of A3G catalytic activity and processivity that has been previously characterized [193,194].

To investigate further if the A3F 231V was partially resistant to Vif-mediated degradation we used V5-tagged A3F 231V and A3F 231I constructs. The tags were used to avoid any possible recognition differences by the native antibody, which recognizes unknown epitopes in the C-terminal half, and to facilitate quantification of immunoblots at lower plasmid transfection

amounts. We repeated the experiment in Figure 3.1D using 50 ng of V5-tagged A3F expression plasmids in the presence or absence of Vif during a single-cycle replication experiment. The V5-tagged A3F-231V had the same features as the untagged enzyme and was expressed 5-fold more than the V5-tagged A3F-231I and was partially resistant to Vif-mediated degradation, in contrast to V5-tagged A3F-231I (Figure 3.2A). The higher steady state levels of the A3F 231V also resulted in higher levels of virion encapsidation in comparison to the A3F 231I, in the presence and absence of Vif (Figure 3.2A). Having established that the higher steady state levels of A3F 231V were indeed due to protein levels in cells and not differences in antibody recognition, we used the V5-tagged constructs to address the basis for the A3F 231V partial protection from Vif-mediated degradation. We transfected different plasmid amounts of the V5-tagged A3F variants that resulted in approximately equal steady state cellular expression. Using 10 ng of A3F 231V-V5 plasmid and 50 ng of A3F 231I-V5 plasmid, we cotransfected increasing amounts of Vif from an expression vector. Quantification of the immunoblot demonstrated that at these transfection amounts, in the absence of Vif, the steady state levels of both A3F 231V-V5 and A3F 231I-V5 were similar (Figure 3.2B). With increasing amounts of transfected Vif, the induced degradation of both A3F variants were also similar (Figure 3.2B-C). Thus, the partial protection from Vif observed in Figure 3.2A and Figure 3.1C-E was not due to an inherent biochemical property of the A3F 231V, but rather the data suggest that the partial protection from Vif-mediated degradation was due to a higher steady state level of A3F 231V that exceeded Vif degradation capacity. Consistent with this line of reasoning, we observed stability differences between A3F 231V-V5 and A3F 231I-V5 after cyclohexamide treatment of transfected cells (50 ng plasmid transfected for each A3F variant). The A3F 231V steady state protein levels decreased 2-fold 10 h after cycloheximide treatment. In contrast, the A3F 231I steady state protein levels decreased 2-fold within 2 h after cycloheximide treatment and continued to decrease through the time course with an ultimate 4-fold decrease in protein level (figure 3.2D-E).

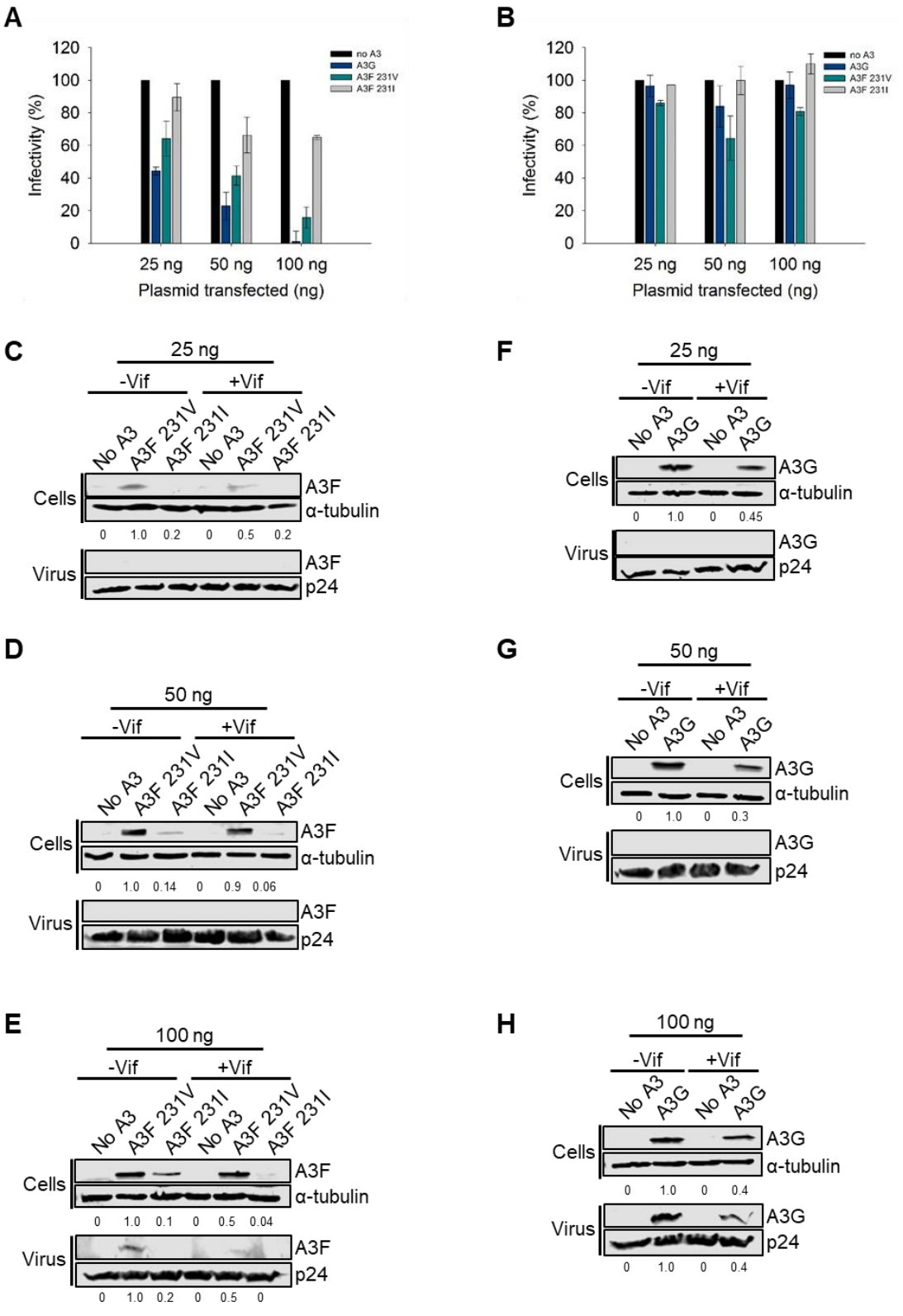

A
Infectivity (%)
120
100
80
60
40
20
0
no A3
A3G
A3F 231V
A3F 231I
25 ng
50 ng
100 ng
Plasmid transfected (ng)
B
Infectivity (%)
120
100
80
60
40
20
0
no A3
A3G
A3F 231V
A3F 231I
25 ng
50 ng
100 ng
Plasmid transfected (ng)
C
25 ng
-Vif
+Vif
No A3
A3F 231V
A3F 231I
No A3
A3F 231V
A3F 231I
Cells
A3F
α-tubulin
0 1.0 0.2 0 0.5 0.2
Virus
A3F
p24
F
25 ng
-Vif
+Vif
No A3
A3G
No A3
A3G
Cells
A3G
α-tubulin
0 1.0 0 0.45
Virus
A3G
p24
D
50 ng
-Vif
+Vif
No A3
A3F 231V
A3F 231I
No A3
A3F 231V
A3F 231I
Cells
A3F
α-tubulin
0 1.0 0.14 0 0.9 0.06
Virus
A3F
p24
G
50 ng
-Vif
+Vif
No A3
A3G
No A3
A3G
Cells
A3G
α-tubulin
0 1.0 0 0.3
Virus
A3G
p24
E
100 ng
-Vif
+Vif
No A3
A3F 231V
A3F 231I
No A3
A3F 231V
A3F 231I
Cells
A3F
α-tubulin
0 1.0 0.1 0 0.5 0.04
Virus
A3F
p24
0 1.0 0.2 0 0.5 0
H
100 ng
-Vif
+Vif
No A3
A3G
No A3
A3G
Cells
A3G
α-tubulin
0 1.0 0 0.4
Virus
A3G
p24
0 1.0 0 0.4

Figure 3.1 A3F 231V and A3F 231I have different HIV-1 restriction abilities. infectivity was measured by β-galactosidase activity in reporter cells infected with HIV-1 ΔVif that was produced in the absence or presence of untagged A3G, A3F 231V or A3F 231I. Results were normalized to the no A3 condition. (B) HIV-1 +Vif infectivity was measured by β-galactosidase activity in reporter cells infected with HIV +Vif that was produced in the absence or presence of untagged A3G, A3F231V or A3F231I. Results were normalized to the no A3 condition. (A-B) Error bars represent the standard deviations of the mean calculated from three independent experiments. (C-H) Immunoblotting with A3F native antibody was used to detect (C) 25 ng (D) 50 ng and (E) 100 ng transfected A3F expressed in (C-E) cells and (E) encapsidated into HIV-1 virions in -Vif and +Vif conditions. Immunoblotting with A3G native antibody was used to detect (F) 25 ng (G) 50 ng and (H) 100 ng transfected A3G expressed in (F-H) cells and (H) encapsidated into HIV-1 virions in -Vif and +Vif conditions. The cell lysate and virion loading controls were α-tubulin and p24, respectively. (C-H) One representative blot from three independent experiments is shown. The A3 expression levels shown below the blots were calculated by setting a –Vif condition to 1.0 and determining the relative values of other lanes. Immunoblots were conducted from three independent experiments and a representative blot is shown.

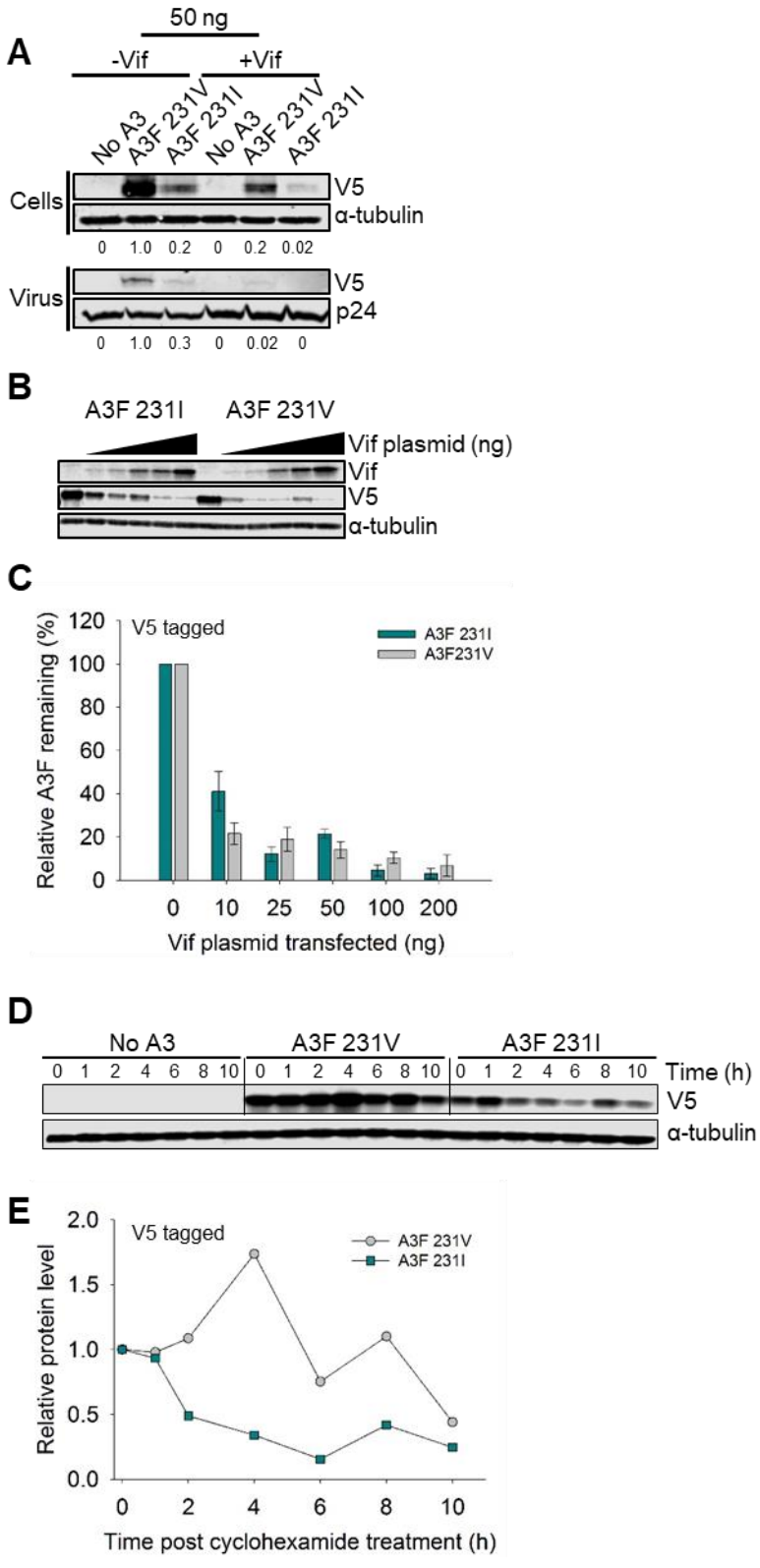

A
50 ng
-Vif
+Vif
No A3
A3F 231V
A3F 231I
No A3
A3F 231V
A3F 231I
Cells
V5
α-tubulin
0 1.0 0.2 0 0.2 0.02
Virus
V5
p24
0 1.0 0.3 0 0.02 0
B
A3F 231I
A3F 231V
Vif plasmid (ng)
Vif
V5
α-tubulin
C
V5 tagged
A3F 231I
A3F231V
Relative A3F remaining (%)
120
100
80
60
40
20
0
0 10 25 50 100 200
Vif plasmid transfected (ng)
D
No A3
A3F 231V
A3F 231I
0 1 2 4 6 8 10 0 1 2 4 6 8 10 0 1 2 4 6 8 10
Time (h)
V5
α-tubulin
E
V5 tagged
A3F 231V
A3F 231I
Relative protein level
2.0
1.5
1.0
0.5
0.0
0 2 4 6 8 10
Time post cyclohexamide treatment (h)

Figure 3.2 A3F 231V is expressed at higher steady state levels in cells. (A) Immunoblotting with V5 antibody after 50 ng of expression plasmid was transfected into 293T cells in the presence of HIV-1 to detect A3F 231V-V5 and A3F 231I-V5 in cells and encapsidated into HIV-1 virions in -Vif and +Vif conditions. (B) A3F 231V-V5 and A3F 231I-V5 plasmids were transfected into 293T cells with amounts that resulted in equal steady state expression levels. Increasing amounts of Vif were cotransfected. The amount of Vif-mediated degradation was detected by analysing the intensity of the bands detected with antibody to the V5 tag. One representative blot from three independent experiments is shown. (C) Quantified intensities of the bands from immunoblotting are shown. Error bars represent the standard deviations of the mean calculated from three independent experiments. (D) Immunoblotting with V5 antibody after 50 ng of expression plasmid was transfected into 293T cells to detect A3F 231V-V5 and A3F 231I-V5 after cycloheximide treatment. (E) Quantification of the protein levels after cycloheximide treatment showed differences in the stability of A3F 231V-V5 and A3F 231I-V5. Immunoblots were conducted from three independent experiments and a representative blot is shown.

3.4.2 A3F 231V induces more mutations than A3F 231I in HIV-1 proviral DNA

We PCR amplified a 351 bp region of HIV-1 *protease* for sequencing under the conditions shown in Figure 3.1A (untagged A3s, 50 ng transfection amount) to determine the G→A mutations induced by A3F 231V and A3F 231I. We observed that in the absence of Vif that G→A mutations per kb were 2.46 for A3F 231V and 1.42 for A3F 231I (Table 3.1). The mutations were primarily in the A3F mutation context (Table 3.1, GA→AA). This ~2-fold difference in mutations is consistent with the decrease in infectivity (Figure 3.1A, 50 ng). However, both A3F variants were still less able to induce mutations than A3G (Table 3.1, 13.49 G→A mutations/kb). This at least 5-fold difference in A3F and A3G induced mutations was consistent with previous data and has been found to be due to differences in enzyme processivity [133]. As a result, these data emphasize the importance of the inherent biochemical characteristics in restriction capacity between A3s. However, when comparing A3F variants, the level of encapsidation appeared to be a determining factor that could increase A3F restriction efficiency for HIV-1 (Table 3.1, Figure 3.1E and Figure 3.2A).

Importantly in the presence of Vif, the number of A3-induced mutations by A3F 231V and A3G were less disparate. A3G did not show significant restriction in the presence of Vif and accordingly induced 36-fold less mutations than in the absence of Vif (Table 3.1, 0.37 G→A mutations/kb). Consistent with the ability to partially avoid Vif-mediated degradation, the A3F 231V only had a 4.8-fold decrease in induced mutations (Table 3.1, 0.51 G→A mutations/kb). Thus, in the presence of Vif the A3F 231V induced mutations were 1.4-fold more than A3G (Table 3.1). In the presence of Vif, the A3F 231I induced mutations decreased 11-fold to 0.13 G→A mutations/kb, resulting in 2.3-fold less mutations than A3G and 4-fold less mutations than A3F 231V (Table 3.1).

3.4.3 A3F genotypes are commonly heterozygous

Since the A3F genotypes showed different abilities to restrict HIV and different steady state expression levels, we were interested to determine how often in the human population these different variant alleles were heterozygous or homozygous. This was of interest since we have previously shown that A3G and A3F can hetero-oligomerize in the absence of RNA [144]. If the A3F variants were expressed in the same cell and hetero-oligomerized, the A3F expression levels and restriction ability could potentially be altered.

We obtained genotype data for the corresponding SNP (rs2076101) that results in the 231V/I polymorphism from Ensembl (release 94) [267]. The cumulative allele data from people of African, Mixed American, East Asian, European and South Asian populations showed that there was an equal frequency of the GTC (G) allele that resulted in the 231V polymorphism and ATC (A) allele that resulted in the 231I polymorphism (Table 3.2). Per population, the allelic frequencies varied, with only the European (EUR) population having an approximately equal frequency of the G and A alleles (Table 3.2). The Mixed American (AMR) and South Asian (SAS) were similar having approximately a 0.40 frequency of the G allele (Table 3.2). The African (AFR) and East Asian (EAS) populations were biased for one allele, with AFR populations having a 0.81 frequency of the G and EAS having a 0.71 frequency of the A (Table 3.2). Despite these differences in allelic frequencies, there were high frequencies across all populations of the heterozygous genotype (A/G). The average frequency across all populations for the heterozygous genotype was 0.42 and was similar to the individual genotype frequencies of the EUR, AMR, EAS, and SAS populations (Table 3.2). Even the AFR population with the lowest allele frequency of the G (0.04), had an A/G genotype frequency of 0.30 (Table 3.2). These data demonstrate that for many populations, the heterozygotic genotype is dominant. The exceptions are the AFR population where the homozygous G/G genotype is dominant (0.66) and the EAS population where the A/A genotype is dominant (0.50) (Table 3.2).

Table 3.1 Analysis of A3-induced mutagenesis in HIV-1 proviral DNA.

Virus Condition	A3 enzyme	Base pairs sequenced	G→A mutations (total)	GG→AG mutations (total)	GA→AA mutations (total)	G→A mutations (per kb)	GG→AG mutations (per kb)	GA→AA mutations (per kb)
-Vif	No A3	7722	2	0	0	0.26	0	0
	A3F 231V	7722	19	2	14	2.46	0.26	1.81
	A3F 231I	7020	10	2	6	1.42	0.28	0.85
	A3G	6669	90	83	5	13.4	12.45	0.75
+Vif	No A3	15795	3	0	1	0.19	0	0.06
	A3F 231V	15795	8	0	5	0.51	0	0.32
	A3F 231I	15093	2	0	2	0.13	0	0.13
	A3G	16142	6	4	1	0.37	0.24	0.06

A 351 bp region of HIV-1 *protease* was PCR amplified from the three independent single-cycle replication experiments shown in Figure 3.1. Clones were sequenced, aligned with Clustal Omega[268], and analyzed using Hypermut[269].

3.4.4 A3F 231V and A3F 231I hetero-oligomerize with each other and A3G

Based on these results showing that different A3F variants can be coexpressed within an individual, we determined whether the A3F variants could hetero-oligomerize with each other or A3G. A3F has been shown to form homo-oligomers and to hetero-oligomerize with A3G, in the absence of RNA [144]. To determine oligomerization, we used a cotransfection scheme where one A3 had an HA-tag and the other A3 had a V5-tag and determined if one A3 could co-immunoprecipitate the other A3. To ensure uniform co-expression of A3s on a per cell basis we used the pVIVO2 plasmid that has two transcription units within a single vector (see Materials and Methods) [144]. We used the V5-tagged A3 to immunoprecipitate the binding partner, in the presence of RNase A, and found through blotting for HA that A3G-HA can interact with A3F 231I-V5, as shown previously (Figure 3.3) [144]. We also determined that A3G-HA can interact with A3F 231V-V5 and that A3F 231V-HA can interact with A3F 231I-V5 (Figure 3.3).

3.4.5 A3F 231V promotes steay state levels of A3F321I and A3G

Using the same pVIVO2 expression vectors that were used in the co-immunoprecipitations we conducted single-cycle replication assays to assess if combined expression changed the HIV-1 restriction capability of the coexpressing A3s compared to the individual enzymes. To facilitate immunblotting of the samples resulting from the infectivity assays, we transfected tagged versions of the A3 enzymes (50 ng) using a combination of HA- and V5- tags so that each A3 was discernable by immunoblotting. The A3G was always HA-tagged, A3F 231I was always V5-tagged and the A3F 231V was HA- or V5- tagged depended on whether it was coexpressed with A3G or A3F. In the absence of Vif, the A3G-HA alone and A3G-HA/A3F 231V-V5 decreased the infectivity similarly (Figure 3.4A). The A3G-HA/A3F 231I-V5 coexpression also decreased HIV infectivity, but the restriction was 3-fold less than the A3G-HA/A3F 231V-V5 restriction, although this was not a statistically significant difference (Figure 3.4A). However, both A3F 231I-V5 and A3F 231V-V5 alone restricted the replication of HIV ΔVif 3-fold ($p \leq 0.001$) and 6-fold less ($p \leq 0.001$), respectively, than when in combination with A3G (Figure 3.4A). These data extend previous data that demonstrated enhanced restriction of HIV when A3G and A3F are coexpressed [126,144,237]. Notably, the V5-tagged versions of A3F 231I and A3F 231V were able to decrease HIV infectivity more than the untagged proteins (Figure 3.1A and Figure 3.4A), but the difference in restriction between the two A3F variants was the same (Figure 3.1A, 50 ng, 1.6-fold, $p \leq 0.01$;

Figure 3.4A, 50 ng, 1.5-fold, $p \leq 0.001$). We also coexpressed A3F 231I-V5 and A3F 231V-HA (A3F 231I/A3F 231V) and found that the decrease in HIV infectivity was 2-fold ($p \leq 0.05$) and 1.5-fold ($p \leq 0.001$) more than each A3F alone, respectively. These data suggest that the weaker restriction activity of A3F 231I does not negatively affect the stronger restriction activity of A3F 231V. In addition, in the presence of Vif, coexpressed A3F 231V/A3F 231I was able to partially restrict HIV-1 (Figure 3.4A, $p \leq 0.01$).

Immunoblotting of cell and virion lysates from the infectivity experiments demonstrated that the A3F 231V resulted in increased steady state levels and encapsidation of coexpressed proteins, A3F 231I or A3G. Since each coexpression condition had an HA- and V5- tagged A3, samples from the same cell lysate had to be detected on separate V5 (Figure 3.4B) or HA (Figure 3.4C) blots. In both the presence and absence of Vif, increased HIV restriction correlated with increased steady state expression of A3F 231I-V5 when coexpressed with A3F 231V-HA (Figure 3.4B). In the absence of Vif and the presence of A3F 231V-HA, the A3F 231I-V5 expression increased 8.4-fold (Figure 3.4B). In the presence of Vif and A3F 231V-HA, the A3F 231I-V5 expression increased 2-fold (Figure 3.4B). A3G-HA expression levels in cells were increased in the presence of A3F 231V-V5 by 1.5-fold in the absence of Vif and 5-fold in the presence of Vif (Figure 3.4C). Effects on encapsidation were also observed in the presence of A3F 231V. Consistent with the partial restriction ability of A3F 231V/A3F 231I in the presence of Vif (Figure 3.4A, 66% infectivity), there was encapsidation of coexpressed A3F 231I-V5 (Figure 3.4B, detectable encapsidation) and A3F 231V-HA (Figure 3.4C, 5-fold more in the presence of Vif) even in the presence of Vif. Although A3G-HA encapsidation in the presence (3-fold) and absence (5-fold) of Vif was increased when coexpressed with A3F 231V-V5, the HIV-1 infectivity in the presence of Vif was not decreased more than A3G-HA alone (Figure 3.4A and Figure 3.4C), consistent with previous results that Vif inhibits encapsidated A3G activity [193,194]. Altogether, the data indicate that when A3F 231V is expressed, it can increase steady state levels and encapsidation of A3F 231I and A3G (Figure 3.4B and Figure 3.4C). Although there was no enhancement of A3G-HA encapsidation levels in the presence of A3F 231I-V5, A3G-HA was able to enhance A3F 231I-V5 cellular expression 10-fold in the absence of Vif and 8-fold in the presence of Vif (Figure 3.4B). The presence of A3G-HA resulted in higher A3F 231I-V5 virion encapsidation levels in the absence of Vif (10-fold), but no change in the presence of Vif (Figure 3.4B, A3F 231I-V5 not detectable). Altogether, these data demonstrate that different A3s can

influence cellular expression levels and virion encapsidation, with A3F 231V having the highest effects.

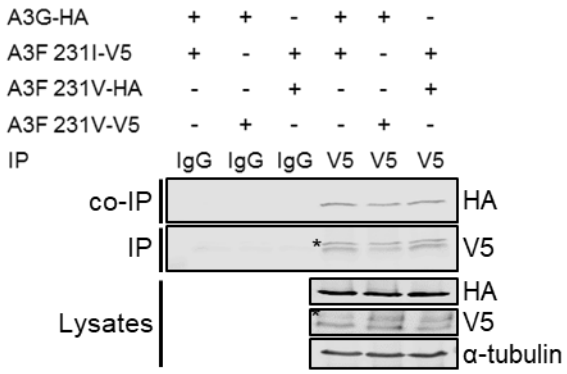


Figure 3.3 A3F 231V and A3F 231I commonly occur as a heterozygous genotype and can interact in cells. The immunoprecipitation from cell lysates used either anti-V5 antibody or rabbit IgG (mock). Co-immunoprecipitation of A3F 231V-HA with A3F 231I-V5 or A3G-HA with A3F 231V-V5 and A3F 231I-V5 was detected through the HA tag. The presence of the V5-tagged protein in the co-immunoprecipitation samples is also shown. The experimental and mock samples were from the same lysate. The lysate blot demonstrates the cellular expression of α-tubulin, HA, and V5. A nonspecific band in the V5 blot is marked with an asterisk.

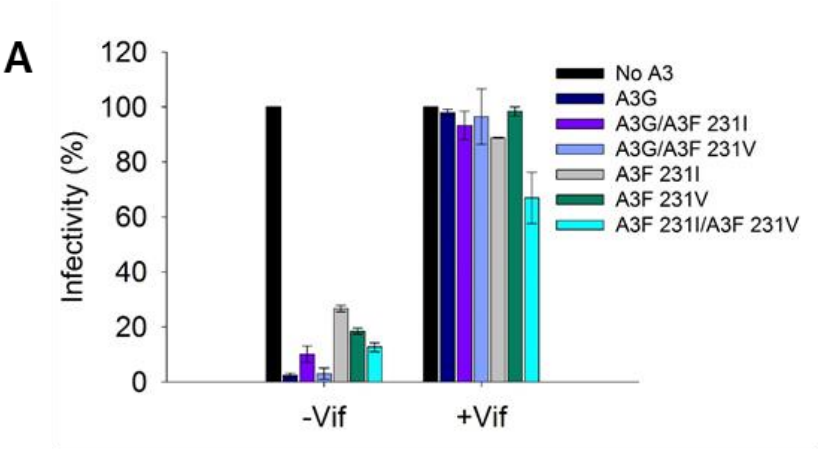
Table 3.2 Population analysis of alleles and genotype of A3F SNP rs2076101 that results in A3F 231V/I.

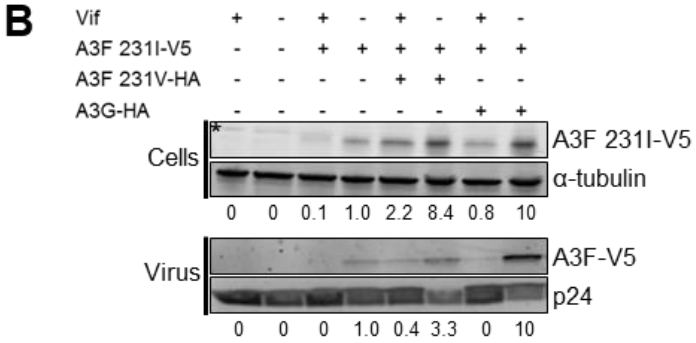
Population[1]	Allele frequency (G/A)[2]	Genotype frequency
ALL	0.50/0.50	G/G: 0.29 A/A: 0.28 A/G: 0.42
AFR	0.81/0.19	G/G: 0.66 A/A: 0.04 A/G: 0.30
EUR	0.51/0.49	G/G: 0.29 A/A: 0.27 A/G: 0.45
AMR	0.38/0.62	G/G: 0.14 A/A: 0.38 A/G: 0.49
EAS	0.29/0.71	G/G: 0.07 A/A: 0.50 A/G: 0.43
SAS	0.39/0.61	G/G: 0.14 A/A: 0.35 A/G: 0.51

Data were obtained from Ensembl genome browser (release 94) [267].

[1] Populations are abbreviated as All (ALL), African (AFR), European (EUR), Mixed American (AMR), East Asian (EAS), and South Asian (SAS).

[2] The A3F SNP rs2076101 is a G to A change that codes for A3F 231V (G) or A3F 231I (A).


A
Infectivity (%)
120
100
80
60
40
20
0
-Vif
+Vif
No A3
A3G
A3G/A3F 231I
A3G/A3F 231V
A3F 231I
A3F 231V
A3F 231I/A3F 231V


B
Vif
A3F 231I-V5
A3F 231V-HA
A3G-HA
Cells
A3F 231I-V5
α-tubulin
0 0 0.1 1.0 2.2 8.4 0.8 10
Virus
A3F-V5
p24
0 0 0 1.0 0.4 3.3 0 10

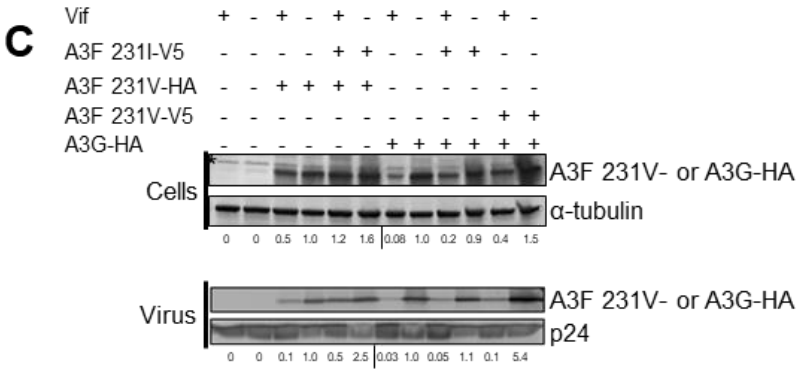

C
Vif
A3F 231I-V5
A3F 231V-HA
A3F 231V-V5
A3G-HA
Cells
A3F 231V- or A3G-HA
α-tubulin
0 0 0.5 1.0 1.2 1.6 0.08 1.0 0.2 0.9 0.4 1.5
Virus
A3F 231V- or A3G-HA
p24
0 0 0.1 1.0 0.5 2.5 0.03 1.0 0.05 1.1 0.1 5.4

Figure 3.4 Co-expression of A3F 231V and A3F 231I results in enhanced HIV-1 restriction ability. (A) HIV-1 -Vif and +Vif infectivity was measured by β-galactosidase activity in reporter cells infected with HIV-1 that was produced in the absence or presence of A3G-HA, A3F 231V-V5, A3F 231V-HA, or A3F 231I-V5 in the combinations shown in the graph legend. Error bars represent the standard deviations of the mean calculated from three independent experiments. (B) Immunoblotting with V5 antibody was used to detect cellular expression and virus encapsidation of A3F 231I-V5 in the presence of A3F 231V-HA or A3G-HA. (C) Immunoblotting with HA antibody was used to detect cellular expression and virus encapsidation of A3F 231V-HA or A3G-HA in the presence of A3F 231I-V5 or A3F 231V-V5. (B-C) Samples in these panels are from the same cell lysates. One representative blot from three independent experiments is shown. The A3 expression levels shown below blots were calculated by setting a –Vif condition for A3F 231I, A3F 231V, or A3G to 1.0 and determining the relative values of other lanes. A nonspecific band in the V5 and HA blots was observed and is marked with an asterisk. Immunoblots were conducted from three independent experiments and a representative blot is shown.

3.4.7 In the presence and absence of Vif, coexpressed A3F 231V and A3F 231I induce more mutations in HIV-1 proviral DNA

To determine if the increased encapsidation of A3 enzymes resulted in increases in A3-induced mutations, we PCR amplified and sequenced a 351 bp portion of HIV-1 *protease*. The data demonstrated that the tagged versions of A3F 231V, A3F 231I, and A3G induced similar numbers of mutations to their untagged versions (+Vif and –Vif, Table 3.1 and Table 3.3). When the tagged A3F 231V and A3F 231I are coexpressed and encapsidate into ΔVif HIV-1 virons, there is a 2-fold increase in mutations from A3F 231I alone and a 1.4-fold increase in mutations for A3F 231V alone (Table 3.3). Based on analysis of encapsidation, this is likely due to increased encapsidation of A3F 231V (Figure 3.5D). The conditions where A3G was coexpressed with an A3F, in HIV-1 ΔVif producer cells, did not have significant increases in the numbers of mutations above A3G alone, but when we examined the ratio of A3G induced mutations to A3F induced mutations based on the mutational footprint, we observed a 2.6 fold (A3F 231V) and 2-fold (A3F 231I) increase in A3F-induced mutations in the presence of A3G (Table 3.3). Of more significance is that in the presence of Vif, there is a 2-fold increase in mutations from the A3F 231V alone condition (Table 3.3). Since A3F 231V and A3F 231I have the same deamination motif, we could not discern which variant is more active. However, based on analysis of encapsidation, since both had higher levels in virions (Figure 3.5C-D), the deaminations could be due to both A3F variants.

Table 3.3 Analysis of combined A3 expression on the induced mutagenesis in HIV-1 proviral DNA.

Virus Condition	A3 enzyme	Base pairs sequenced	G→A mutations (total)	GG→AG mutations (total)	GA→AA mutations (total)	G→A mutations (per kb)	GG→AG mutations (per kb)	GA→AA mutations (per kb)
-Vif	No A3	10881	0	0	0	0	0	0
	A3F 231V-V5	10530	27	1	26	2.56	0.09	2.47
	A3F 231I-V5	8073	15	1	14	1.86	0.12	1.73
	A3F 231V-HA/A3F231I-V5	12285	45	1	44	3.66	0.08	3.58
	A3G	6669	121	109	11	18.1	16.3	1.64
	A3G-HA/A3F 231V-V5	10530	172	102	70	16.3	9.69	6.64
	A3G-HA/ A3F 231I-V5	12987	134	95	39	10.3	7.32	3.00
+Vif	No A3	11934	0	0	0	0	0	0
	A3F 231V-V5	10530	6	1	5	0.57	0.09	0.47
	A3F 231I-V5	10881	1	0	1	0.09	0	0.09
	A3F 231V-HA/A3F231I-V5	11232	11	0	11	0.98	0	0.98

A 351 bp region of HIV-1 *protease* was PCR amplified from the three independent single-cycle replication experiments shown in Figure 3.4. Clones were sequenced, aligned with Clustal Omega [268], and analyzed using Hypermut [269].

3.5 Discussion

In HIV-1 genomes isolated from HIV+ individuals, there are mutational footprints in the (+) DNA at both 5′GG→5′AG and 5′GA→5′AA motifs [127,156,195-204,258]. The 5′GG→5′AG are clearly attributable to A3G [125]. The origins of the 5′GA→5′AA mutations have been attributed to stable A3H haplotype (II, V, VII) induced mutations, due to an inability of some HIV-1 Vif variants to efficiently induce degradation of A3H [127,264]. The role of A3D in inducing these mutations is known to be minor, but still significant [131,136,218,236,270]. However, the role of A3F has been significant in some reports, but not others [126,131,133,217,218,233,234]. In individual experiments with A3F, the mutational load is lower than A3G and A3H and this seemed incongruent with data demonstrating that A3F exerts a direct selective force on HIV-1 Vif and provides a protective effect in pretreatment HIV-1 cohorts [77,133,135,160,217,271]. Here we provide a possible explanation for these data demonstrating that the A3F 231V in comparison to the A3F 231I is more stable in cells, is partially protected from Vif-mediated degradation, is more highly encapsidated into virions, induces more mutations, and can increase encapsidation of coexpressed A3F 231I and A3G. These data suggest that the A3F 231V through its own action and hetero-oligomerization with A3F 231I or A3G is a contributor to G→A mutations in HIV+ individuals.

A protective effect of A3F 231V was demonstrated through genetic association analysis of pretreatment cohorts [160]. Through this analysis it was shown that even when data were adjusted for known genetic factors (*HLA* allele and CCR5) the presence of even one A3F 231V allele could significantly delay progression to clinical AIDS and lower viral load [160]. The authors suggested that this was due to partial resistance of A3F 231V to Vif, which had also been reported previously [126,160]. Here we show that the A3F 231V is protected from Vif mediated degradation in cells more than A3F 231I and this results in more A3F 231V encapsidation (Figure 3.1). However, the A3F 231V was not “resistant” to Vif mediated degradation, consistent with the 231 amino acid not being within the Vif binding interface [186,190,265,266]. The protection from Vif appears to be due to the higher steady state level of A3F 231V in cells due to enhanced stability (Figure 3.1 and Figure 3.2). Due to the high frequency of heterozygosity of the A3F 231V and A3F 231I and the lack of sensitive native A3F antibodies, we were unable to demonstrate the higher expression of A3F 231V in primary cells (Table 3.2 and data not shown). However, when equal amounts of A3F 231V and A3F 231I were expressed in cells, the capacity of Vif to induce degradation of both

enzymes was similar (Figure 3.2B-C). Rather, the data suggest that the higher steady state levels of A3F in cells exceeds the degradation capacity of Vif. Despite the protective effect of A3F 231V in HIV+ individuals [160], the partial protection of A3F 231V from Vif mediated degradation resulted in at most a 20% decrease in infectivity in a Vif+ HIV-1 single-cycle replication experiment (Figure 3.1B). More consistent with the An *et al*. study was 4-fold more G→A mutations induced by A3F 231V than A3F 231I in the presence of Vif (Table 3.1) [160]. These small effects in our cell-based studies are similar to data available for the A3G H186R polymorphism. The A3G H186R polymorphism results in only a 2-fold decrease in A3G processivity and mutations induced in an *in vitro* assay [134], but is nonetheless linked with faster progression to AIDS in HIV+ individuals [210,211]. Altogether, the data suggest that these small differences identified in cell culture can have an effect in HIV+ individuals. A possible explanation for this is that the small effects observed in cell-based studies become cumulative in multiple HIV replication cycles. In addition or alternatively, the coexpression and hetero-oligomerization of A3 enzymes may amplify this effect. The coexpression of A3F 231V and A3F 231I increased the encapsidation of both A3F variants in the presence of Vif and resulted in a corresponding 1.7-fold increase in mutations and infectivity of 66% (Figure 3.4B, Figure 3.4C, and Table 3.3). Our data and previous analysis demonstrate that there are synergies between A3 enzymes that can enhance anti-HIV activity, either directly due to changes in the biochemical characteristics of the enzymes or due to enhanced encapsidation in the presence of Vif (Figure 3.4) [144]. Despite observations that A3F 231V on its own is partially protected from Vif-mediated degradation and our observation that A3F 231V appears to protect A3G from Vif-mediated degradation, the mechanism is not known [126,160].

Further, the reason for the stability differences between A3F 231V and A3F 231I remain to be elucidated. Consistent with our observations, a nucleotide variant prediction model suggested that the A3F linked SNPs can have an effect on stability (Figure 3.1 and Figure 3.2) [160]. The amino acid 108 is on predicted helix 3 in the noncatalytic N-terminal domain (NTD) and the amino acid 231 is on β-strand 2 in the catalytic C-terminal domain [92,186,272,273]. Neither of these structures are near the catalytic center, predicted nucleic acid binding motifs, or the Vif interface [92,186,190,272,273]. Although V231I is a conservative amino acid change, a Val to Ile amino acid change in other proteins, e.g., transthyretin, has been found to destabilize equilibrium kinetics of oligomerization and result in physiological dysfunction [274]. The A108S is a nonpolar to polar amino acid change and the Ser can form hydrogen bonds, unlike Ala, which may alter protein stability [275].

3.4.6 Conclusions

In aim 2 we observed that the virus that was exposed to multiple rounds of deamination had less RT activity and subsequently less viability compared to the no A3 or singly exposed to A3s conditions. Moreover, our experiments showed that there were also variations within the virus populations that were exposed to multiple rounds of deaminations. Since the type of the A3 enzyme that was used in the experiments were similar, we hypothesized that this variation could be due to polymorphism in A3 family. Genotyping of the CEM and PBMCs that endogenously express A3s showed that our donor's PBMCs harbored the less restrictive variant of A3F with Isoleucine in position 231 and CEM cells were heterozygous for A3F (A3F231I/A3F231V). Considering the results of other studies showing that harboring A3F 231V allele was associated with delayed progression to AIDS [160] and the results of aim 2 that A3s cooperate to insure maximum restriction of HIV-1 we sought to investigate the impact of A3F polymorphism on HIV-1 infectivity in aim3.

Our data demonstrated that the A3F 231V has higher steady state expression levels in cells and virion encapsidation than A3F 231I, in the presence and absence of Vif. This results in larger decreases in HIV-1 infectivity and higher mutation rates. The disruption of Vif-mediated degradation through hetero-oligomerization of A3s may be an example of increasing host restriction factor diversity to combat virus evolution [164,276]. Altogether, this supports a model in which not only the genetic diversity of the host for each A3, but the ability of these A3s to functionally interact can contribute to a diverse and multifaceted protective response during HIV-1 infection.

3.6 Materials and Methods

3.6.1 Plasmids and Transfection Conditions

To express two A3 enzymes on a single-cell level, the pVIVO2 plasmid (Invivogen) that contains two transcriptional units was used. For comparison to double expression experiments, pVIVO2 was also used to express single A3 enzymes. The cloning of A3G-HA, A3F 108S/231I-V5, and A3G-HA/A3F 108S/231I-V5 into pVIVO2 was previously described[144]. The pcDNA A3F 108A/231V was kindly provided by Reuben Harris (University of Minnesota) and subcloned into pVIVO2 using the same cloning strategy as previously reported to create A3F 108A/231V-V5, A3F 108A/231V-HA, and A3G-HA/A3F 108A/231V-V5, and A3F 108A/231V-HA/A3F

108S/231I-V5[144]. Primer sequences are available on request. Untagged constructs of A3G, A3F 108S/231I, and A3F 108A/231V were expressed from pcDNA and were previously reported[126,133]. Since the A3F polymorphisms at 108 and 231 are in strong linkage disequilibrium, for brevity we refered to the A3F forms as either 231I or 231V[160,212,233].

3.6.2 Single-cycle infectivity assays

VSV-G pseudotyped HIV-1 LAI ΔVif ΔEnv were produced by transfecting 1 x 10^5 293T cells per well of a 12-well plate. The 293T cells were maintained in DMEM with 10% FBS. GeneJuice (Novagen) transfection reagent was used as per manufacturer's protocol. Cells were transfected with 500 ng of pHIV-1 LAI ΔVif ΔEnv, 180 ng of pMDG, which expresses VSV-G, and 25, 50, or 100 ng of A3 expression plasmid[260,277]. The following pVIVO2 constructs were transfected for expression of tagged A3s: A3G-HA, A3F 231V-V5, A3F 231I-V5, A3G-HA/A3F 231I-V5, A3G-HA/A3F 231V-V5, and A3F 231V-HA/A3F 231I-V5. The following pcDNA3 constructs were transfected for expression of untagged A3s: A3G, A3F 231V, and A3F 231I. To equalize the amount of plasmid DNA transfected, empty pVIVO2 or pcDNA3 vectors were used. After 24 h the media was changed and virus containing supernatants were harvested 24 h after the media change. Supernatants were filtered through a 0.45 μm polyvinylidene difluoride (PVDF) syringe filter.

For infection of a reporter cell line to determine infectivity 1 x 10^4 cells per well of a 96-well plate containing either HeLa CD4+ HIV-1 LTR-β-gal (MAGI) or TZM-Bl cells were infected with a serial dilution of virus in the presence of 8 μg/mL polybrene. Forty-eight hours after infection the cells were washed with PBS and infectivity was measured through colorimetric detection using a β-galactosidase assay reagent (Pierce) and spectrophotometer. Infectivity of each virus was compared by using the infectivity of the No A3 condition as 100%.

3.6.3 Quantitative immunoblotting

293T cells expressing No A3, A3G-HA, A3F 231V-V5, A3F 231I-V5, A3F-231V-HA, A3G-HA/A3F 231I-V5, A3G-HA/A3F 231V-V5, and A3F 231V-HA/A3F 231I-V5 from the single-cycle infectivity assays were detected using rabbit anti-HA (1:1000, Sigma) or rabbit anti-V5 (1:500, Sigma). 293T cells expressing No A3 or untagged A3G, A3F 231V, or A3F 231I were detected with ApoC17 rabbit antiserum (1:1000; Cat# 10082, NIH AIDS Reagent Program) for

A3G or rabbit anti-APOBEC3F antibody, C-term (1:500; Cat# GTX47211, Genetex) for A3F[278,279]. A3s were detected in cell lysates and virions. Cells were lysed using 2x Laemmli Buffer and 40 μg total protein was used. Virus was concentrated using Retro-X (Clontech) following the manufacturer's protocol and 20 μL of concentrated virus was used. Loading controls for cell lysates (α-tubulin, Sigma) and virus (p24, Cat #3537, NIH AIDS Reagent Program) were detected with mouse monoclonal antibodies. Secondary detection was performed using Licor IRDye antibodies produced in goat (IRDye 680-labeled anti-rabbit and IRDye 800-labeled anti mouse).

To detect Vif-mediated degradation of A3s, 1 x 10^5 293T cells per well of a 24-well plate were transfected with 50 ng of A3F 231I-V5, 10 ng A3F 231V-V5, or empty pVIVO2 (No A3) and a titration of Vif_{LAI} expression plasmid (0,10, 25, 50, 100, and 200 ng). A 5-fold greater amount of A3F 231I was required to equalize the expression of both A3Fs. GeneJuice (Novagen) transfection reagent was used as described by the manufacturer. The media (DMEM and 10% FBS) was changed 24 h after the transfection. The cells were harvested 24 h after the media change with 2x Laemmli buffer and 40 μg total protein was used. Vif was detected with HIV-1 Vif monoclonal antibody (1:1000; Cat# 6459, NIH AIDS Reagent Program), followed by incubation with secondary goat anti-mouse IRDye 680RD. The V5-tagged A3F and α-tubulin was detected as described.

3.6.4 Co-immunoprecipitation assay

The co-immunoprecipitation method was previously described[144,262]. In brief, 2.5 x 10^6 293T cells per 75 cm^2 flask were transfected with 1 μg total DNA using GeneJuice (Novagen) as per manufacturer instructions. The plasmids transfected expressed A3G-HA/A3F 231I-V5, A3G-HA/A3F 231V-V5 or A3F 231V-HA/A3F-231I-V5. Each cell lysate was split into two fractions that were either used in a co-immunoprecipitation with nonspecific mouse IgG (mock) or the mouse anti-V5 (experiment). The co-immunoprecipitations were conducted in the presence of RNase A[144]. A mouse anti-V5 (1:1000, Sigma) was used for immunoprecipitations and proteins were detected on the membrane using the same V5 antibody, mouse anti-HA (1:1000, Sigma), and polyclonal rabbit anti-α-tubulin (1:1000; Sigma). Secondary antibodies used were Licor IRDye antibodies IRDye 800-labeled goat anti-mouse and 680-labeled goat anti-rabbit secondary antibody.

3.6.5 Sequencing of integrated proviral DNA

For proviral sequencing, 1 x 10^5 293T cells per well of a 24-well plate were infected with supernatant containing virus in the presence of 8 μg/mL polybrene. The plates were spinoculated at 800 x *g* for 1 h. Cell were harvested after 48 h by removing the media, washing with PBS, and lysing the cells and extracting DNA with DNAzol (Invitrogen). The PCR amplification of the *protease* region of HIV-1 (351 bp) and treatment of DNA with DpnI was carried out as previously described[133]. Sequences were analyzed with Clustal Omega[268] and Hypermut[269].

3.6.6 Genotyping Donor Peripheral Blood Mononuclear Cells

PBMCs were isolated from whole blood samples using Ficoll-Paque (GE-Healthcare) density gradient centrifugation in Sepmate50 tubes (STEMCELL Technologies). DNA was isolated from donor PBMCs using a PAXgene Blood DNA Kit (Qiagen). The A3G and A3F polymorphic sites were amplified using PCR. The A3G H186R polymorphism was detected using the following primers: (forward) 5′ AATTTCAGCACTGTTGGAGC and (reverse) 5′ GAGAATCTCCCCCAGCAT. The A3F V231I polymorphism was detected using the following primers: (forward) 5′ AGCCTATGGTCGGAACGAAA and (reverse) 5′ CTGGTTTCGGAAGACGCC. The PCR amplification used Q5 High Fidelity polymerase (New England Biolabs) and cycling conditions were 95° C for 5 min followed by 45 cycles of 95° C for 10 sec, 52° C for 20 sec, 72 ° C for 20 sec. Amplicons were cloned into pJET1.2 using the CloneJET PCR cloning kit (ThermoScientific). Amplicons were Sanger sequenced at the National Research Council of Canada (Saskatoon, Saskatchewan). The resulting A3G and A3F sequences were translated using the Expasy Translate tool to identify polymorphisms [280]. Six clones from each donor were analyzed for A3G H186R and A3F I231V polymorphisms to determine homozygosity or heterozygosity of each allele in combination with allelic frequency.

3.7 Ethics Statement

This study has been reviewed and approved by the Research and Ethics Board at the University of Saskatchewan (Bio# 14-62 and 13-293). Peripheral Blood Mononuclear Cells (PBMCs) were isolated from de-identified individuals of European ancestry. In accordance with OCAP® principles and in consultation with several Indigenous stakeholders, no samples of Indigenous

Peoples were included. Donors were either HIV+ and recruited through the Positive Living Program (PLP) at the Royal University Hospital, Saskatoon or HIV- Donors (HD).

3.8 Acknowledgements

This research was supported by a Canadian Institutes for Health Research Grant MOP-137090. The authors thank Ayan Moallin and Stephen Patrick for technical assistance. The following reagents were obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: anti-ApoC17 from Dr. Klaus Strebel (Cat# 10082) and Vif monoclonal antibody (Cat# 6459) from Dr. Michael H. Malim.

4.0. GENERAL DISCUSSION AND FUTURE DIRECTIONS

HIV-1 is a major health problem that has taken over 35 million lives. HIV/AIDS results in deficiency of immune system by attacking CD4+T cells causing rare cancers and opportunistic infections for an infected individual. Constant genetic diversity in the replicating population of the virus is a challenging obstacle facing developing a successful HIV vaccine. Understanding all the contributors to this genetic diversity helps scientists to tackle this barrier and design successful antiretroviral drugs. A3 enzymes despite their intrinsic antiretroviral activity are one of the potential contributors to this genetic variation. However, there have been controversial reports on this contribution and the extent to which A3-mediated mutagenesis can take part in the development of drug resistance is not clear. Despite A3G being discovered 17 years ago, we are still discovering new functions of these enzymes. In this work we incorporated recent findings about A3 enzymes to come to a broader conclusion regarding the contribution of A3 enzymes to virus evolution.

In 2017 a study from our lab by Ara *et al*. demonstrated that A3F and A3G form a tetrameric hetero-oligomer when expressed together. Using purified proteins and size exclusion chromatography, the study showed that the hetero-oligomer contained a trimeric A3F and a monomer A3G. The co-IP results further showed that this hetero-oligomerization is through an RNA independent protein-protein interaction in the cells[144]. This hetero-oligomerization resulted in cooperative restriction of ΔVif HIV-1. This enhanced antiretroviral activity in the absence of Vif was due to improved mutagenesis of the A3s and the ability of the tetramer to slow down the reverse transcriptase by acting as a road block[144]. However, the extent to which these proteins

oligomerize in the presence of a WT HIV and their impact on the replication of WT HIV was unknown. Here, we showed that in the presence of Vif, A3F indeed hetero-oligomerizes with both A3G and A3F. This hetero-oligomerization equips the A3F with a partial protection to the action of Vif. This resistance recovers the mutagenic activity of A3F in the presence of Vif which would have less than the error rate of RT if A3F was expressed alone. However, the reason behind this partial protection is not fully understood.

Similar results were found with coexpressioin of A3F 231I/231V. The enhanced HIV restriction activity of the hetero-oligomer of A3F231I/231V appears to be due to the better expression level of A3F231V that results first, in the presence of increased numbers of available A3F molecules in the cytoplasm which can exceed its degradative capacity and second, to increase that stability of the hetero-oligomerized partner (A3F 231I). For the A3F/G, the protection from Vif could be due to the conformational changes that oligomerization offers which can block the Vif binding site on the A3s. However, Vif is known to disrupt A3G oligomerization which can affect the processivity of the enzyme[262]. The improved protection of only A3F in this tetrameric hetero-oligomer suggests that Vif is unable to break up A3G hetero-oligomerization with A3F.

This work demonstrates that A3F is more important for HIV restriction than previously thought, either alone or in combination with A3G in restriction of HIV. As a result, this work suggests that more polymorphisms should be taken into account when the effect of A3s in retroviral suppression and evolution is being studied since the properties of A3s are different when hetero-oligomerized or alone. This raises the question about the accountability of the results of the studies overlooking the effect of the A3s when hetero-oligomerized. The CEM cells with endogenous expression of A3s are extensively used in HIV field and have heterozygous A3F alleles. This heterozygosity provides different properties for the cells than the homozygosity by A3F 231I; this could explain different conclusions of A3F involvement in HIV restriction from experiments from different labs that did or did not use CEM cells. Besides CEM cells, other A3 producing cells like H9 and SupT1are widely used and have not been genotyped.

In this study we focused on A3F and A3G since these proteins are the most restrictive of the five members and are commonly expressed together in the human population[230,231]. However, based on the fact that A3s get encapsidated in a non-competitive way through binding nonspecifically to cellular RNAs that HIV Gag is also bound to, the potential hetero-

oligomerization of different A3 pairs should be considered as this may affect activity. For example, A3D is the least active enzyme[137], but is most closely related to A3F, suggesting that they may be able to hetero-oligomerize. The effect of this on HIV restriction activity should be investigated to determine if hetero-oliomerization always promotes HIV restriction, or if some A3 pairs result in less A3 restriction. Further, in our experiments we used lab isolates of virus had a limited duration of infection (30 days), but in the future it would be interesting to use clinical isolates of viruses from early and chronic infections and continually passage the virus to determine how quickly the A3-mediated mutations accumulate in proviral genomes.

Overall, this study showed that A3 proteins cooperate to enhance the viral suppression. The data support a model that this cooperation is a counteraction against Vif-mediated degradation to increase the efficiency of the HIV-1 restriction. Perhaps this mechanism and affinity of A3s to hetero-oligomerize could be used develop more efficient antiviral drugs.

5.0. REFERENCES

1 Tebit, D. M. & Arts, E. J. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. *Lancet Infect Dis* **11**, 45-56, doi:10.1016/S1473-3099(10)70186-9 (2011).

2 de Sousa, J. D., Muller, V., Lemey, P. & Vandamme, A. M. High GUD incidence in the early 20 century created a particularly permissive time window for the origin and initial spread of epidemic HIV strains. *PLoS One* **5**, e9936, doi:10.1371/journal.pone.0009936 (2010).

3 Robertson, D. L. *et al.* HIV-1 nomenclature proposal. *Science* **288**, 55-56 (2000).

4 Buonaguro, L., Tornesello, M. L. & Buonaguro, F. M. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *J Virol* **81**, 10209-10219, doi:10.1128/JVI.00872-07 (2007).

5 Campbell-Yesufu, O. T. & Gandhi, R. T. Update on human immunodeficiency virus (HIV)-2 infection. *Clin Infect Dis* **52**, 780-787, doi:10.1093/cid/ciq248 (2011).

6 Seitz, R. Human Immunodeficiency Virus (HIV). *German Advisory Committee Blood, Subgroup* **Assessment of Pathogens Transmissibe by Blood** (2016).

7 O'Donovan, D. *et al.* Maternal plasma viral RNA levels determine marked differences in mother-to-child transmission rates of HIV-1 and HIV-2 in The Gambia. MRC/Gambia Government/University College London Medical School working group on mother-child transmission of HIV. *AIDS* **14**, 441-448 (2000).

8 Okoye, A. A. & Picker, L. J. CD4(+) T-cell depletion in HIV infection: mechanisms of immunological failure. *Immunol Rev* **254**, 54-64, doi:10.1111/imr.12066 (2013).

9 Coffin, J., Hughes, S. H. & E., V. H. Retroviruses. *Cold Spring Harbor Laboratory Press* (1997).

10 Florom-Smith, A. L. & De Santis, J. P. Exploring the concept of HIV-related stigma. *Nurs Forum* **47**, 153-165, doi:10.1111/j.1744-6198.2011.00235.x (2012).

11 Herek, G. M. & Glunt, E. K. An epidemic of stigma. Public reactions to AIDS. *Am Psychol* **43**, 886-891 (1988).

12 Mervis, R. J. *et al.* The gag gene products of human immunodeficiency virus type 1: alignment within the gag open reading frame, identification of posttranslational modifications, and evidence for alternative gag precursors. *J Virol* **62**, 3993-4002 (1988).

13 Accola, M. A., Hoglund, S. & Gottlinger, H. G. A putative alpha-helical structure which overlaps the capsid-p2 boundary in the human immunodeficiency virus type 1 Gag precursor is crucial for viral particle assembly. *J Virol* **72**, 2072-2078 (1998).

14 Wyatt, R. & Sodroski, J. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* **280**, 1884-1888 (1998).

15 Barre-Sinoussi, F., Ross, A. L. & Delfraissy, J. F. Past, present and future: 30 years of HIV research. *Nat Rev Microbiol* **11**, 877-883, doi:10.1038/nrmicro3132 (2013).

16 Lu, K., Heng, X. & Summers, M. F. Structural determinants and mechanism of HIV-1 genome packaging. *J Mol Biol* **410**, 609-633, doi:10.1016/j.jmb.2011.04.029 (2011).

17 Harris, R. S., Hultquist, J. F. & Evans, D. T. The restriction factors of human immunodeficiency virus. *J Biol Chem* **287**, 40875-40883, doi:10.1074/jbc.R112.416925 (2012).

18 Selby, M. J., Bain, E. S., Luciw, P. A. & Peterlin, B. M. Structure, sequence, and position of the stem-loop in tar determine transcriptional elongation by tat through the HIV-1 long terminal repeat. *Genes Dev* **3**, 547-558 (1989).
19 Cochrane, A. W., Perkins, A. & Rosen, C. A. Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: relevance of nucleolar localization to function. *J Virol* **64**, 881-885 (1990).
20 Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W. & Luhrmann, R. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**, 475-483 (1995).
21 Collins, D. R. & Collins, K. L. HIV-1 accessory proteins adapt cellular adaptors to facilitate immune evasion. *PLoS Pathog* **10**, e1003851, doi:10.1371/journal.ppat.1003851 (2014).
22 Strebel, K. HIV accessory proteins versus host restriction factors. *Curr Opin Virol* **3**, 692-699, doi:10.1016/j.coviro.2013.08.004 (2013).
23 Sheehy, A. M., Gaddis, N. C., Choi, J. D. & Malim, M. H. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* **418**, 646-650, doi:10.1038/nature00939 (2002).
24 Sova, P. & Volsky, D. J. Efficiency of viral DNA synthesis during infection of permissive and nonpermissive cells with vif-negative human immunodeficiency virus type 1. *J Virol* **67**, 6322-6326 (1993).
25 Harris, R. S. *et al.* DNA deamination mediates innate immunity to retroviral infection. *Cell* **113**, 803-809 (2003).
26 Mariani, R. *et al.* Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. *Cell* **114**, 21-31 (2003).
27 Jager, S. *et al.* Vif hijacks CBF-beta to degrade APOBEC3G and promote HIV-1 infection. *Nature* **481**, 371-375, doi:10.1038/nature10693 (2012).
28 Mehle, A. *et al.* Vif overcomes the innate antiviral activity of APOBEC3G by promoting its degradation in the ubiquitin-proteasome pathway. *J Biol Chem* **279**, 7792-7798, doi:10.1074/jbc.M313093200 (2004).
29 Yu, X. *et al.* Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* **302**, 1056-1060, doi:10.1126/science.1089591 (2003).
30 Sheehy, A. M., Gaddis, N. C. & Malim, M. H. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nature medicine* **9**, 1404-1407, doi:10.1038/nm945 (2003).
31 Conticello, S. G., Harris, R. S. & Neuberger, M. S. The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. *Curr Biol* **13**, 2009-2013 (2003).
32 Stopak, K., de Noronha, C., Yonemoto, W. & Greene, W. C. HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. *Molecular cell* **12**, 591-601 (2003).
33 Guo, Y. *et al.* Structural basis for hijacking CBF-beta and CUL5 E3 ligase complex by HIV-1 Vif. *Nature* **505**, 229-233, doi:10.1038/nature12884 (2014).
34 Kirchhoff, F., Greenough, T. C., Brettler, D. B., Sullivan, J. L. & Desrosiers, R. C. Brief report: absence of intact nef sequences in a long-term survivor with nonprogressive HIV-1 infection. *N Engl J Med* **332**, 228-232, doi:10.1056/NEJM199501263320405 (1995).

35 Salvi, R. *et al.* Grossly defective nef gene sequences in a human immunodeficiency virus type 1-seropositive long-term nonprogressor. *J Virol* **72**, 3646-3657 (1998).
36 Catucci, M., Venturi, G., Romano, L., Valensin, P. E. & Zazzi, M. Analysis of the HIV-1 nef gene in five intravenous drug users with long-term nonprogressive HIV-1 infection in Italy. *J Med Virol* **60**, 294-299 (2000).
37 Welker, R., Harris, M., Cardel, B. & Krausslich, H. G. Virion incorporation of human immunodeficiency virus type 1 Nef is mediated by a bipartite membrane-targeting signal: analysis of its role in enhancement of viral infectivity. *J Virol* **72**, 8833-8840 (1998).
38 Forshey, B. M. & Aiken, C. Disassembly of human immunodeficiency virus type 1 cores in vitro reveals association of Nef with the subviral ribonucleoprotein complex. *J Virol* **77**, 4409-4414 (2003).
39 Collins, K. L., Chen, B. K., Kalams, S. A., Walker, B. D. & Baltimore, D. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. *Nature* **391**, 397-401, doi:10.1038/34929 (1998).
40 Pawlak, E. N. & Dikeakos, J. D. HIV-1 Nef: a master manipulator of the membrane trafficking machinery mediating immune evasion. *Biochim Biophys Acta* **1850**, 733-741, doi:10.1016/j.bbagen.2015.01.003 (2015).
41 Pawlak, E. N., Dirk, B. S., Jacob, R. A., Johnson, A. L. & Dikeakos, J. D. The HIV-1 accessory proteins Nef and Vpu downregulate total and cell surface CD28 in CD4(+) T cells. *Retrovirology* **15**, 6, doi:10.1186/s12977-018-0388-3 (2018).
42 Rosa, A. *et al.* HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation. *Nature* **526**, 212-217, doi:10.1038/nature15399 (2015).
43 Usami, Y., Wu, Y. & Gottlinger, H. G. SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. *Nature* **526**, 218-223, doi:10.1038/nature15400 (2015).
44 Andrew, A. & Strebel, K. HIV-1 Vpu targets cell surface markers CD4 and BST-2 through distinct mechanisms. *Mol Aspects Med* **31**, 407-417, doi:10.1016/j.mam.2010.08.002 (2010).
45 Ewart, G. D., Sutherland, T., Gage, P. W. & Cox, G. B. The Vpu protein of human immunodeficiency virus type 1 forms cation-selective ion channels. *J Virol* **70**, 7108-7115 (1996).
46 Schubert, U. *et al.* Identification of an ion channel activity of the Vpu transmembrane domain and its involvement in the regulation of virus release from HIV-1-infected cells. *FEBS Lett* **398**, 12-18 (1996).
47 Neil, S. J., Zang, T. & Bieniasz, P. D. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451**, 425-430, doi:10.1038/nature06553 (2008).
48 Perez-Caballero, D. *et al.* Tetherin inhibits HIV-1 release by directly tethering virions to cells. *Cell* **139**, 499-511, doi:10.1016/j.cell.2009.08.039 (2009).
49 Janvier, K. *et al.* The ESCRT-0 component HRS is required for HIV-1 Vpu-mediated BST-2/tetherin down-regulation. *PLoS Pathog* **7**, e1001265, doi:10.1371/journal.ppat.1001265 (2011).
50 Jia, B. *et al.* Species-specific activity of SIV Nef and HIV-1 Vpu in overcoming restriction by tetherin/BST2. *PLoS Pathog* **5**, e1000429, doi:10.1371/journal.ppat.1000429 (2009).

51 Bour, S., Schubert, U., Peden, K. & Strebel, K. The envelope glycoprotein of human immunodeficiency virus type 2 enhances viral particle release: a Vpu-like factor? *J Virol* **70**, 820-829 (1996).
52 Tungaturthi, P. K. *et al.* Role of HIV-1 Vpr in AIDS pathogenesis: relevance and implications of intravirion, intracellular and free Vpr. *Biomed Pharmacother* **57**, 20-24 (2003).
53 Sherman, M. P. *et al.* HIV-1 Vpr displays natural protein-transducing properties: implications for viral pathogenesis. *Virology* **302**, 95-105 (2002).
54 Di Marzio, P., Choe, S., Ebright, M., Knoblauch, R. & Landau, N. R. Mutational analysis of cell cycle arrest, nuclear localization and virion packaging of human immunodeficiency virus type 1 Vpr. *J Virol* **69**, 7909-7916 (1995).
55 Laguette, N. *et al.* Premature activation of the SLX4 complex by Vpr promotes G2/M arrest and escape from innate immune sensing. *Cell* **156**, 134-145, doi:10.1016/j.cell.2013.12.011 (2014).
56 Jowett, J. B. *et al.* The human immunodeficiency virus type 1 vpr gene arrests infected T cells in the G2 + M phase of the cell cycle. *J Virol* **69**, 6304-6313 (1995).
57 Fekairi, S. *et al.* Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell* **138**, 78-89, doi:10.1016/j.cell.2009.06.029 (2009).
58 Munoz, I. M. *et al.* Coordination of structure-specific nucleases by human SLX4/BTBD12 is required for DNA repair. *Mol Cell* **35**, 116-127, doi:10.1016/j.molcel.2009.06.020 (2009).
59 Blondot, M. L., Dragin, L., Lahouassa, H. & Margottin-Goguet, F. How SLX4 cuts through the mystery of HIV-1 Vpr-mediated cell cycle arrest. *Retrovirology* **11**, 117, doi:10.1186/s12977-014-0117-5 (2014).
60 Tristem, M., Marshall, C., Karpas, A., Petrik, J. & Hill, F. Origin of vpx in lentiviruses. *Nature* **347**, 341-342, doi:10.1038/347341b0 (1990).
61 Hrecka, K. *et al.* Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* **474**, 658-661, doi:10.1038/nature10195 (2011).
62 Berger, A. *et al.* SAMHD1-deficient CD14+ cells from individuals with Aicardi-Goutieres syndrome are highly susceptible to HIV-1 infection. *PLoS Pathog* **7**, e1002425, doi:10.1371/journal.ppat.1002425 (2011).
63 Kim, B., Nguyen, L. A., Daddacha, W. & Hollenbaugh, J. A. Tight interplay among SAMHD1 protein level, cellular dNTP levels, and HIV-1 proviral DNA synthesis kinetics in human primary monocyte-derived macrophages. *J Biol Chem* **287**, 21570-21574, doi:10.1074/jbc.C112.374843 (2012).
64 Lahouassa, H. *et al.* SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat Immunol* **13**, 223-228, doi:10.1038/ni.2236 (2012).
65 Engelman, A. & Cherepanov, P. The structural biology of HIV-1: mechanistic and therapeutic insights. *Nat Rev Microbiol* **10**, 279-290, doi:10.1038/nrmicro2747 (2012).
66 Le Grice, S. F. Human immunodeficiency virus reverse transcriptase: 25 years of research, drug discovery, and promise. *J Biol Chem* **287**, 40850-40857, doi:10.1074/jbc.R112.389056 (2012).
67 Jacobo-Molina, A. & Arnold, E. HIV reverse transcriptase structure-function relationships. *Biochemistry* **30**, 6351-6356 (1991).

68 Rausch, J. W. & Le Grice, S. F. 'Binding, bending and bonding': polypurine tract-primed initiation of plus-strand DNA synthesis in human immunodeficiency virus. *Int J Biochem Cell Biol* **36**, 1752-1766, doi:10.1016/j.biocel.2004.02.016 (2004).
69 Hu, W. S. & Hughes, S. H. HIV-1 reverse transcription. *Cold Spring Harb Perspect Med* **2**, doi:10.1101/cshperspect.a006882 (2012).
70 Suzuki, Y. & Craigie, R. The road to chromatin - nuclear entry of retroviruses. *Nat Rev Microbiol* **5**, 187-196, doi:10.1038/nrmicro1579 (2007).
71 Aiken, C. Viral and cellular factors that regulate HIV-1 uncoating. *Curr Opin HIV AIDS* **1**, 194-199, doi:10.1097/01.COH.0000221591.11294.c1 (2006).
72 Karn, J. & Stoltzfus, C. M. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb Perspect Med* **2**, a006916, doi:10.1101/cshperspect.a006916 (2012).
73 Sundquist, W. I. & Krausslich, H. G. HIV-1 assembly, budding, and maturation. *Cold Spring Harb Perspect Med* **2**, a006924, doi:10.1101/cshperspect.a006924 (2012).
74 Carlson, L. A. *et al.* Three-dimensional analysis of budding sites and released virus suggests a revised model for HIV-1 morphogenesis. *Cell Host Microbe* **4**, 592-599, doi:10.1016/j.chom.2008.10.013 (2008).
75 Morita, E. & Sundquist, W. I. Retrovirus budding. *Annu Rev Cell Dev Biol* **20**, 395-425, doi:10.1146/annurev.cellbio.20.010403.102350 (2004).
76 Pettit, S. C. *et al.* The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. *J Virol* **68**, 8017-8027 (1994).
77 Adolph, M. B., Love, R. P. & Chelico, L. Biochemical Basis of APOBEC3 Deoxycytidine Deaminase Activity on Diverse DNA Substrates. *ACS Infect Dis* **4**, 224-238, doi:10.1021/acsinfecdis.7b00221 (2018).
78 WHO, U., UNAIDS. Global update on HIV treatmnt 2013: results, impact and opportunities. *Geneva: WHO* (2013).
79 WHO, G. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. (2016).
80 UNAIDS, G. How AIDS changes everything - MDG6: 15 years, 15 lessons of hope from the AIDS response. (2015).
81 Cohen, M. S. *et al.* Prevention of HIV-1 infection with early antiretroviral therapy. *N Engl J Med* **365**, 493-505, doi:10.1056/NEJMoa1105243 (2011).
82 Weller, I. V. & Williams, I. G. ABC of AIDS. Antiretroviral drugs. *BMJ* **322**, 1410-1412 (2001).
83 WHO, G. Policy Brief: Consolidated guidelines on HIV prevention, diagnosis, treatment and care for key populations, 2016

update. (2017).
84 Webcast, C. Narrowing the gap in lafe expectancy for HIV+ compared with HIV-indivduals... Life expextancy gap narrows but persists. *Conference on Retroviruses and opportunistic infections (CROI)* (2016).
85 Carr, A. *et al.* A syndrome of peripheral lipodystrophy, hyperlipidaemia and insulin resistance in patients receiving HIV protease inhibitors. *AIDS* **12**, F51-58 (1998).
86 Shen, L. *et al.* Dose-response curve slope sets class-specific limits on inhibitory potential of anti-HIV drugs. *Nat Med* **14**, 762-766, doi:10.1038/nm1777 (2008).

87 Pinti, M., Salomoni, P. & Cossarizza, A. Anti-HIV drugs and the mitochondria. *Biochim Biophys Acta* **1757**, 700-707, doi:10.1016/j.bbabio.2006.05.001 (2006).
88 WHO, G. Global report on early warning indicators of HIV drug resistance: technical report. (2017).
89 WHO, G. Global action plan on HIV drug resistance 2017–2021. (2017).
90 Abram, M. E., Ferris, A. L., Shao, W., Alvord, W. G. & Hughes, S. H. Nature, position, and frequency of mutations made in a single cycle of HIV-1 replication. *J Virol* **84**, 9864-9878, doi:10.1128/JVI.00915-10 (2010).
91 HIV/AIDS data and statistics. (WHO).
92 Feng, Y., Baig, T. T., Love, R. P. & Chelico, L. Suppression of APOBEC3-mediated restriction of HIV-1 by Vif. *Front Microbiol* **5**, 450, doi:10.3389/fmicb.2014.00450 (2014).
93 Conticello, S. G. The AID/APOBEC family of nucleic acid mutators. *Genome Biol* **9**, 229, doi:10.1186/gb-2008-9-6-229 (2008).
94 Salter, J. D., Bennett, R. P. & Smith, H. C. The APOBEC Protein Family: United by Structure, Divergent in Function. *Trends Biochem Sci* **41**, 578-594, doi:10.1016/j.tibs.2016.05.001 (2016).
95 Sharma, S. *et al.* APOBEC3A cytidine deaminase induces RNA editing in monocytes and macrophages. *Nat Commun* **6**, 6881, doi:10.1038/ncomms7881 (2015).
96 Sharma, S., Patnaik, S. K., Taggart, R. T. & Baysal, B. E. The double-domain cytidine deaminase APOBEC3G is a cellular site-specific RNA editing enzyme. *Sci Rep* **6**, 39100, doi:10.1038/srep39100 (2016).
97 Chelico, L., Pham, P., Petruska, J. & Goodman, M. F. Biochemical basis of immunological and retroviral responses to DNA-targeted cytosine deamination by activation-induced cytidine deaminase and APOBEC3G. *J Biol Chem* **284**, 27761-27765, doi:10.1074/jbc.R109.052449 (2009).
98 Navaratnam, N. *et al.* The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *J Biol Chem* **268**, 20709-20712 (1993).
99 Teng, B., Burant, C. F. & Davidson, N. O. Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science* **260**, 1816-1819 (1993).
100 Ohtsubo, H. *et al.* APOBEC2 negatively regulates myoblast differentiation in muscle regeneration. *Int J Biochem Cell Biol* **85**, 91-101, doi:10.1016/j.biocel.2017.02.005 (2017).
101 Sato, Y. *et al.* Deficiency in APOBEC2 leads to a shift in muscle fiber type, diminished body mass, and myopathy. *J Biol Chem* **285**, 7111-7118, doi:10.1074/jbc.M109.052977 (2010).
102 Rogozin, I. B., Basu, M. K., Jordan, I. K., Pavlov, Y. I. & Koonin, E. V. APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases predicted by computational analysis. *Cell Cycle* **4**, 1281-1285, doi:10.4161/cc.4.9.1994 (2005).
103 Koito, A. & Ikeda, T. Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases. *Front Microbiol* **4**, 28, doi:10.3389/fmicb.2013.00028 (2013).
104 Jarmuz, A. *et al.* An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. *Genomics* **79**, 285-296, doi:10.1006/geno.2002.6718 (2002).

105 Chiu, Y. L. & Greene, W. C. The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu Rev Immunol* **26**, 317-353, doi:10.1146/annurev.immunol.26.021607.090350 (2008).
106 Conticello, S. G., Thomas, C. J., Petersen-Mahrt, S. K. & Neuberger, M. S. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol Biol Evol* **22**, 367-377, doi:10.1093/molbev/msi026 (2005).
107 LaRue, R. S. *et al.* Guidelines for naming nonprimate APOBEC3 genes and proteins. *J Virol* **83**, 494-497, doi:10.1128/JVI.01976-08 (2009).
108 LaRue, R. S. *et al.* The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. *BMC Mol Biol* **9**, 104, doi:10.1186/1471-2199-9-104 (2008).
109 Navarro, F. *et al.* Complementary function of the two catalytic domains of APOBEC3G. *Virology* **333**, 374-386, doi:10.1016/j.virol.2005.01.011 (2005).
110 Wittkopp, C. J., Adolph, M. B., Wu, L. I., Chelico, L. & Emerman, M. A Single Nucleotide Polymorphism in Human APOBEC3C Enhances Restriction of Lentiviruses. *PLoS Pathog* **12**, e1005865, doi:10.1371/journal.ppat.1005865 (2016).
111 OhAinle, M., Kerns, J. A., Li, M. M., Malik, H. S. & Emerman, M. Antiretroelement activity of APOBEC3H was lost twice in recent human evolution. *Cell Host Microbe* **4**, 249-259, doi:10.1016/j.chom.2008.07.005 (2008).
112 Harari, A., Ooms, M., Mulder, L. C. & Simon, V. Polymorphisms and splice variants influence the antiretroviral activity of human APOBEC3H. *J Virol* **83**, 295-303, doi:10.1128/JVI.01665-08 (2009).
113 Wang, X. *et al.* Analysis of human APOBEC3H haplotypes and anti-human immunodeficiency virus type 1 activity. *J Virol* **85**, 3142-3152, doi:10.1128/JVI.02049-10 (2011).
114 Starrett, G. J. *et al.* The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. *Nat Commun* **7**, 12918, doi:10.1038/ncomms12918 (2016).
115 Jern, P. & Coffin, J. M. Host-retrovirus arms race: trimming the budget. *Cell Host Microbe* **4**, 196-197, doi:10.1016/j.chom.2008.08.008 (2008).
116 Jacques, D. A. *et al.* HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis. *Nature* **536**, 349-353, doi:10.1038/nature19098 (2016).
117 York, A., Kutluay, S. B., Errando, M. & Bieniasz, P. D. The RNA Binding Specificity of Human APOBEC3 Proteins Resembles That of HIV-1 Nucleocapsid. *PLoS Pathog* **12**, e1005833, doi:10.1371/journal.ppat.1005833 (2016).
118 Duggal, N. K. & Emerman, M. Evolutionary conflicts between viruses and restriction factors shape immunity. *Nat Rev Immunol* **12**, 687-695, doi:10.1038/nri3295 (2012).
119 Simon, V. *et al.* Natural variation in Vif: differential impact on APOBEC3G/3F and a potential role in HIV-1 diversification. *PLoS Pathog* **1**, e6, doi:10.1371/journal.ppat.0010006 (2005).
120 Fourati, S. *et al.* Partially active HIV-1 Vif alleles facilitate viral escape from specific antiretrovirals. *AIDS* **24**, 2313-2321, doi:10.1097/QAD.0b013e32833e515a (2010).
121 Gabuzda, D. H. *et al.* Role of vif in replication of human immunodeficiency virus type 1 in CD4+ T lymphocytes. *J Virol* **66**, 6489-6495 (1992).

122 Wang, T. *et al.* Distinct viral determinants for the packaging of human cytidine deaminases APOBEC3G and APOBEC3C. *Virology* **377**, 71-79, doi:10.1016/j.virol.2008.04.012 (2008).
123 Yu, Q. *et al.* APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J Biol Chem* **279**, 53379-53386, doi:10.1074/jbc.M408802200 (2004).
124 Yang, B., Chen, K., Zhang, C., Huang, S. & Zhang, H. Virion-associated uracil DNA glycosylase-2 and apurinic/apyrimidinic endonuclease are involved in the degradation of APOBEC3G-edited nascent HIV-1 DNA. *J Biol Chem* **282**, 11667-11675, doi:10.1074/jbc.M606864200 (2007).
125 Yu, Q. *et al.* Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat Struct Mol Biol* **11**, 435-442, doi:10.1038/nsmb758 nsmb758 [pii] (2004).
126 Liddament, M. T., Brown, W. L., Schumacher, A. J. & Harris, R. S. APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo. *Curr Biol* **14**, 1385-1391, doi:10.1016/j.cub.2004.06.050 (2004).
127 Ooms, M. *et al.* HIV-1 Vif adaptation to human APOBEC3H haplotypes. *Cell Host Microbe* **14**, 411-421, doi:10.1016/j.chom.2013.09.006 (2013).
128 Berg, O. G., Winter, R. B. & von Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* **20**, 6929-6948 (1981).
129 von Hippel, P. H. & Berg, O. G. Facilitated target location in biological systems. *J Biol Chem* **264**, 675-678 (1989).
130 Stanford, N. P., Szczelkun, M. D., Marko, J. F. & Halford, S. E. One- and three-dimensional pathways for proteins to reach specific DNA sites. *EMBO J* **19**, 6546-6557, doi:10.1093/emboj/19.23.6546 (2000).
131 Hultquist, J. F. *et al.* Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *J Virol* **85**, 11220-11234, doi:10.1128/JVI.05238-11 (2011).
132 Chelico, L., Pham, P., Calabrese, P. & Goodman, M. F. APOBEC3G DNA deaminase acts processively 3' --> 5' on single-stranded DNA. *Nat Struct Mol Biol* **13**, 392-399, doi:10.1038/nsmb1086 (2006).
133 Ara, A., Love, R. P. & Chelico, L. Different mutagenic potential of HIV-1 restriction factors APOBEC3G and APOBEC3F is determined by distinct single-stranded DNA scanning mechanisms. *PLoS Pathog* **10**, e1004024, doi:10.1371/journal.ppat.1004024 (2014).
134 Feng, Y. & Chelico, L. Intensity of deoxycytidine deamination of HIV-1 proviral DNA by the retroviral restriction factor APOBEC3G is mediated by the noncatalytic domain. *J Biol Chem* **286**, 11415-11426, doi:10.1074/jbc.M110.199604 (2011).
135 Feng, Y. *et al.* Natural Polymorphisms and Oligomerization of Human APOBEC3H Contribute to Single-stranded DNA Scanning Ability. *J Biol Chem* **290**, 27188-27203, doi:10.1074/jbc.M115.666065 (2015).
136 Duggal, N. K., Malik, H. S. & Emerman, M. The breadth of antiviral activity of Apobec3DE in chimpanzees has been driven by positive selection. *J Virol* **85**, 11361-11371, doi:10.1128/JVI.05046-11 (2011).

137 Dang, Y., Wang, X., Esselman, W. J. & Zheng, Y. H. Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. *J Virol* **80**, 10522-10533, doi:10.1128/JVI.01123-06 (2006).
138 Adolph, M. B., Webb, J. & Chelico, L. Retroviral restriction factor APOBEC3G delays the initiation of DNA synthesis by HIV-1 reverse transcriptase. *PLoS One* **8**, e64196, doi:10.1371/journal.pone.0064196 (2013).
139 Iwatani, Y. *et al.* Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G. *Nucleic Acids Res* **35**, 7096-7108, doi:10.1093/nar/gkm750 (2007).
140 Belanger, K., Savoie, M., Rosales Gerpe, M. C., Couture, J. F. & Langlois, M. A. Binding of RNA by APOBEC3G controls deamination-independent restriction of retroviruses. *Nucleic Acids Res* **41**, 7438-7452, doi:10.1093/nar/gkt527 (2013).
141 Bishop, K. N., Verma, M., Kim, E. Y., Wolinsky, S. M. & Malim, M. H. APOBEC3G inhibits elongation of HIV-1 reverse transcripts. *PLoS Pathog* **4**, e1000231, doi:10.1371/journal.ppat.1000231 (2008).
142 Wang, X. *et al.* The cellular antiviral protein APOBEC3G interacts with HIV-1 reverse transcriptase and inhibits its function during viral replication. *J Virol* **86**, 3777-3786, doi:10.1128/JVI.06594-11 (2012).
143 Gillick, K. *et al.* Suppression of HIV-1 infection by APOBEC3 proteins in primary human CD4(+) T cells is associated with inhibition of processive reverse transcription as well as excessive cytidine deamination. *J Virol* **87**, 1508-1517, doi:10.1128/JVI.02587-12 (2013).
144 Ara, A. *et al.* Mechanism of Enhanced HIV Restriction by Virion Coencapsidated Cytidine Deaminases APOBEC3F and APOBEC3G. *J Virol* **91**, doi:10.1128/JVI.02230-16 (2017).
145 Pollpeter, D. *et al.* Deep sequencing of HIV-1 reverse transcripts reveals the multifaceted antiviral functions of APOBEC3G. *Nat Microbiol* **3**, 220-233, doi:10.1038/s41564-017-0063-9 (2018).
146 Kobayashi, T. *et al.* Quantification of deaminase activity-dependent and -independent restriction of HIV-1 replication mediated by APOBEC3F and APOBEC3G through experimental-mathematical investigation. *J Virol* **88**, 5881-5887, doi:10.1128/JVI.00062-14 (2014).
147 Coffin, J. & Swanstrom, R. HIV pathogenesis: dynamics and genetics of viral populations and infected cells. *Cold Spring Harb Perspect Med* **3**, a012526, doi:10.1101/cshperspect.a012526 (2013).
148 Mansky, L. M. & Temin, H. M. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J Virol* **69**, 5087-5094 (1995).
149 Simon-Loriere, E. & Holmes, E. C. Why do RNA viruses recombine? *Nat Rev Microbiol* **9**, 617-626, doi:10.1038/nrmicro2614 (2011).
150 Mulder, L. C., Harari, A. & Simon, V. Cytidine deamination induced HIV-1 drug resistance. *Proc Natl Acad Sci U S A* **105**, 5501-5506, doi:10.1073/pnas.0710190105 (2008).
151 Nowarski, R., Britan-Rosich, E., Shiloach, T. & Kotler, M. Hypermutation by intersegmental transfer of APOBEC3G cytidine deaminase. *Nat Struct Mol Biol* **15**, 1059-1066, doi:10.1038/nsmb.1495 (2008).

152 Delviks-Frankenberry, K. A. *et al.* Minimal Contribution of APOBEC3-Induced G-to-A Hypermutation to HIV-1 Recombination and Genetic Variation. *PLoS Pathog* **12**, e1005646, doi:10.1371/journal.ppat.1005646 (2016).
153 Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M. & Ho, D. D. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271**, 1582-1586 (1996).
154 Chun, T. W. *et al.* Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature* **387**, 183-188, doi:10.1038/387183a0 (1997).
155 Neogi, U. *et al.* Human APOBEC3G-mediated hypermutation is associated with antiretroviral therapy failure in HIV-1 subtype C-infected individuals. *J Int AIDS Soc* **16**, 18472, doi:10.7448/IAS.16.1.18472 (2013).
156 Kieffer, T. L. *et al.* G-->A hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting CD4+ T cells in vivo. *J Virol* **79**, 1975-1980, doi:10.1128/JVI.79.3.1975-1980.2005 (2005).
157 Simen, B. B. *et al.* Low-abundance drug-resistant viral variants in chronically HIV-infected, antiretroviral treatment-naive patients significantly impact treatment outcomes. *J Infect Dis* **199**, 693-701, doi:10.1086/596736 (2009).
158 Li, J. Z. *et al.* Low-frequency HIV-1 drug resistance mutations and risk of NNRTI-based antiretroviral treatment failure: a systematic review and pooled analysis. *JAMA* **305**, 1327-1335, doi:10.1001/jama.2011.375 (2011).
159 Fourati, S. *et al.* Differential impact of APOBEC3-driven mutagenesis on HIV evolution in diverse anatomical compartments. *Aids* **28**, 487-491, doi:10.1097/qad.0000000000000182 (2014).
160 An, P. *et al.* Role of APOBEC3F Gene Variation in HIV-1 Disease Progression and Pneumocystis Pneumonia. *PLoS Genet* **12**, e1005921, doi:10.1371/journal.pgen.1005921 (2016).
161 Simon, V., Bloch, N. & Landau, N. R. Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. *Nat Immunol* **16**, 546-553, doi:10.1038/ni.3156 (2015).
162 Compton, A. A. & Emerman, M. Convergence and divergence in the evolution of the APOBEC3G-Vif interaction reveal ancient origins of simian immunodeficiency viruses. *PLoS Pathog* **9**, e1003135, doi:10.1371/journal.ppat.1003135

PPATHOGENS-D-12-02303 [pii] (2013).
163 Compton, A. A., Hirsch, V. M. & Emerman, M. The host restriction factor APOBEC3G and retroviral Vif protein coevolve due to ongoing genetic conflict. *Cell Host Microbe* **11**, 91-98, doi:S1931-3128(11)00402-1 [pii]

10.1016/j.chom.2011.11.010 (2012).
164 Compton, A. A., Malik, H. S. & Emerman, M. Host gene evolution traces the evolutionary history of ancient primate lentiviruses. *Philos Trans R Soc Lond B Biol Sci* **368**, 20120496, doi:10.1098/rstb.2012.0496 (2013).
165 Mitchell, P. S., Young, J. M., Emerman, M. & Malik, H. S. Evolutionary Analyses Suggest a Function of MxB Immunity Proteins Beyond Lentivirus Restriction. *PLoS Pathog* **11**, e1005304, doi:10.1371/journal.ppat.1005304 (2015).
166 Fregoso, O. I. *et al.* Evolutionary toggling of Vpx/Vpr specificity results in divergent recognition of the restriction factor SAMHD1. *PLoS Pathog* **9**, e1003496, doi:10.1371/journal.ppat.1003496 (2013).

167 Etienne, L., Hahn, B. H., Sharp, P. M., Matsen, F. A. & Emerman, M. Gene loss and adaptation to hominids underlie the ancient origin of HIV-1. *Cell Host Microbe* **14**, 85-92, doi:10.1016/j.chom.2013.06.002 (2013).
168 Harris, R. S. & Anderson, B. D. Evolutionary Paradigms from Ancient and Ongoing Conflicts between the Lentiviral Vif Protein and Mammalian APOBEC3 Enzymes. *PLoS Pathog* **12**, e1005958, doi:10.1371/journal.ppat.1005958 (2016).
169 Nakano, Y. *et al.* A conflict of interest: the evolutionary arms race between mammalian APOBEC3 and lentiviral Vif. *Retrovirology* **14**, 31, doi:10.1186/s12977-017-0355-4 (2017).
170 Sauter, D. & Kirchhoff, F. Multilayered and versatile inhibition of cellular antiviral factors by HIV and SIV accessory proteins. *Cytokine & growth factor reviews* **40**, 3-12, doi:10.1016/j.cytogfr.2018.02.005 (2018).
171 Alteri, C. *et al.* Incomplete APOBEC3G/F Neutralization by HIV-1 Vif Mutants Facilitates the Genetic Evolution from CCR5 to CXCR4 Usage. *Antimicrob Agents Chemother* **59**, 4870-4881, doi:10.1128/aac.00137-15 (2015).
172 Jern, P., Russell, R. A., Pathak, V. K. & Coffin, J. M. Likely role of APOBEC3G-mediated G-to-A mutations in HIV-1 evolution and drug resistance. *PLoS Pathog* **5**, e1000367, doi:10.1371/journal.ppat.1000367 (2009).
173 Casartelli, N. *et al.* The antiviral factor APOBEC3G improves CTL recognition of cultured HIV-infected T cells. *J Exp Med* **207**, 39-49, doi:10.1084/jem.20091933 (2010).
174 Grant, M. & Larijani, M. Evasion of adaptive immunity by HIV through the action of host APOBEC3G/F enzymes. *AIDS Res Ther* **14**, 44, doi:10.1186/s12981-017-0173-8 (2017).
175 Desimmie, B. A. *et al.* Multiple APOBEC3 restriction factors for HIV-1 and one Vif to rule them all. *J Mol Biol* **426**, 1220-1245, doi:10.1016/j.jmb.2013.10.033 (2014).
176 Zhou, X., Evans, S. L., Han, X., Liu, Y. & Yu, X. F. Characterization of the interaction of full-length HIV-1 Vif protein with its key regulator CBFbeta and CRL5 E3 ubiquitin ligase components. *PLoS One* **7**, e33495, doi:10.1371/journal.pone.0033495

PONE-D-12-00989 [pii] (2012).
177 Zhang, W., Du, J., Evans, S. L., Yu, Y. & Yu, X. F. T-cell differentiation factor CBF-beta regulates HIV-1 Vif-mediated evasion of host restriction. *Nature* **481**, 376-379, doi:nature10718 [pii]

10.1038/nature10718 (2012).
178 Stanley, B. J. *et al.* Structural insight into the human immunodeficiency virus Vif SOCS box and its role in human E3 ubiquitin ligase assembly. *J Virol* **82**, 8656-8663, doi:10.1128/jvi.00767-08 (2008).
179 Yu, Y., Xiao, Z., Ehrlich, E. S., Yu, X. & Yu, X. F. Selective assembly of HIV-1 Vif-Cul5-ElonginB-ElonginC E3 ubiquitin ligase complex through a novel SOCS box and upstream cysteines. *Genes Dev* **18**, 2867-2872, doi:10.1101/gad.1250204 (2004).
180 Bergeron, J. R. *et al.* The SOCS-box of HIV-1 Vif interacts with ElonginBC by induced-folding to recruit its Cul5-containing ubiquitin ligase complex. *PLoS Pathog* **6**, e1000925, doi:10.1371/journal.ppat.1000925 (2010).
181 Iwatani, Y. *et al.* HIV-1 Vif-mediated ubiquitination/degradation of APOBEC3G involves four critical lysine residues in its C-terminal domain. *Proc Natl Acad Sci U S A* **106**, 19539-19544, doi:10.1073/pnas.0906652106 (2009).

182 Albin, J. S. *et al.* Dispersed sites of HIV Vif-dependent polyubiquitination in the DNA deaminase APOBEC3F. *J Mol Biol* **425**, 1172-1182, doi:S0022-2836(13)00013-2 [pii]

10.1016/j.jmb.2013.01.010 (2013).
183 Apolonia, L. *et al.* Promiscuous RNA binding ensures effective encapsidation of APOBEC3 proteins by HIV-1. *PLoS Pathog* **11**, e1004609, doi:10.1371/journal.ppat.1004609 (2015).
184 Nowarski, R., Britan-Rosich, E., Shiloach, T. & Kotler, M. Hypermutation by intersegmental transfer of APOBEC3G cytidine deaminase. *Nat Struct Mol Biol* **15**, 1059-1066 (2008).
185 Kitamura, S. *et al.* The APOBEC3C crystal structure and the interface for HIV-1 Vif binding. *Nat Struct Mol Biol* **19**, 1005-1010, doi:10.1038/nsmb.2378 (2012).
186 Nakashima, M. *et al.* Structural Insights into HIV-1 Vif-APOBEC3F Interaction. *J Virol* **90**, 1034-1047, doi:10.1128/jvi.02369-15 (2016).
187 Nakashima, M. *et al.* Mapping Region of Human Restriction Factor APOBEC3H Critical for Interaction with HIV-1 Vif. *J Mol Biol* **429**, 1262-1276, doi:10.1016/j.jmb.2017.03.019 (2017).
188 Ooms, M., Letko, M. & Simon, V. The Structural Interface between HIV-1 Vif and Human APOBEC3H. *J Virol* **91**, doi:10.1128/JVI.02289-16 (2017).
189 Letko, M., Booiman, T., Kootstra, N., Simon, V. & Ooms, M. Identification of the HIV-1 Vif and Human APOBEC3G Protein Interface. *Cell Rep* **13**, 1789-1799, doi:10.1016/j.celrep.2015.10.068 (2015).
190 Richards, C. *et al.* The Binding Interface between Human APOBEC3F and HIV-1 Vif Elucidated by Genetic and Computational Approaches. *Cell Rep* **13**, 1781-1788, doi:10.1016/j.celrep.2015.10.067 (2015).
191 Binning, J. M. *et al.* Fab-based inhibitors reveal ubiquitin independent functions for HIV Vif neutralization of APOBEC3 restriction factors. *PLoS Pathog* **14**, e1006830, doi:10.1371/journal.ppat.1006830 (2018).
192 Opi, S. *et al.* Human immunodeficiency virus type 1 Vif inhibits packaging and antiviral activity of a degradation-resistant APOBEC3G variant. *J Virol* **81**, 8236-8246, doi:JVI.02694-06 [pii]

10.1128/JVI.02694-06 (2007).
193 Britan-Rosich, E., Nowarski, R. & Kotler, M. Multifaceted counter-APOBEC3G mechanisms employed by HIV-1 Vif. *J Mol Biol* **410**, 1065-1076, doi:10.1016/j.jmb.2011.03.058 (2011).
194 Feng, Y., Love, R. P. & Chelico, L. HIV-1 viral infectivity factor (Vif) alters processive single-stranded DNA scanning of the retroviral restriction factor APOBEC3G. *J Biol Chem* **288**, 6083-6094, doi:M112.421875 [pii]

10.1074/jbc.M112.421875 (2013).
195 Janini, M., Rogers, M., Birx, D. R. & McCutchan, F. E. Human immunodeficiency virus type 1 DNA sequences genetically damaged by hypermutation are often abundant in patient peripheral blood mononuclear cells and may be generated during near-simultaneous infection and activation of CD4(+) T cells. *J Virol* **75**, 7973-7986 (2001).
196 Vartanian, J. P., Henry, M. & Wain-Hobson, S. Sustained G-->A hypermutation during reverse transcription of an entire human immunodeficiency virus type 1 strain Vau group O genome. *J Gen Virol* **83**, 801-805, doi:10.1099/0022-1317-83-4-801 (2002).

197 Vartanian, J. P., Meyerhans, A., Asjö, B. & Wain-Hobson, S. Selection, recombination, and G----A hypermutation of human immunodeficiency virus type 1 genomes. *J Virol* **65**, 1779-1788 (1991).
198 Pathak, V. K. & Temin, H. M. Broad spectrum of in vivo forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: substitutions, frameshifts, and hypermutations. *Proc Natl Acad Sci U S A* **87**, 6019-6023 (1990).
199 Bruner, K. M. *et al.* Defective proviruses rapidly accumulate during acute HIV-1 infection. *Nature medicine* **22**, 1043-1049, doi:10.1038/nm.4156 (2016).
200 Pace, C. *et al.* Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with APOBEC3G and vif genetic variation. *J Virol* **80**, 9259-9269, doi:80/18/9259 [pii]

10.1128/JVI.00888-06 (2006).
201 Piantadosi, A., Humes, D., Chohan, B., McClelland, R. S. & Overbaugh, J. Analysis of the percentage of human immunodeficiency virus type 1 sequences that are hypermutated and markers of disease progression in a longitudinal cohort, including one individual with a partially defective Vif. *J Virol* **83**, 7805-7814, doi:10.1128/JVI.00280-09 (2009).
202 de Lima-Stein, M. L. *et al.* In vivo HIV-1 hypermutation and viral loads among antiretroviral-naive Brazilian patients. *AIDS Res Hum Retroviruses* **30**, 867-880, doi:10.1089/aid.2013.0241 (2014).
203 Eyzaguirre, L. M. *et al.* Elevated hypermutation levels in HIV-1 natural viral suppressors. *Virology* **443**, 306-312, doi:10.1016/j.virol.2013.05.019 (2013).
204 Land, A. M. *et al.* Human immunodeficiency virus (HIV) type 1 proviral hypermutation correlates with CD4 count in HIV-infected women from Kenya. *J Virol* **82**, 8172-8182, doi:10.1128/JVI.01115-08 (2008).
205 Harris, R. S. & Dudley, J. P. APOBECs and virus restriction. *Virology* **479-480**, 131-145, doi:10.1016/j.virol.2015.03.012 (2015).
206 Suspene, R., Rusniok, C., Vartanian, J. P. & Wain-Hobson, S. Twin gradients in APOBEC3 edited HIV-1 DNA reflect the dynamics of lentiviral replication. *Nucleic Acids Res* **34**, 4677-4684, doi:gkl555 [pii]

10.1093/nar/gkl555 (2006).
207 Hu, C. *et al.* The HIV-1 central polypurine tract functions as a second line of defense against APOBEC3G/F. *J Virol* **84**, 11981-11993, doi:JVI.00723-10 [pii]

10.1128/JVI.00723-10 (2010).
208 Wurtzer, S. *et al.* Functional central polypurine tract provides downstream protection of the human immunodeficiency virus type 1 genome from editing by APOBEC3G and APOBEC3B. *J Virol* **80**, 3679-3683, doi:80/7/3679 [pii]

10.1128/JVI.80.7.3679-3683.2006 (2006).
209 Adolph, M. B. *et al.* Cytidine deaminase efficiency of the lentiviral viral restriction factor APOBEC3C correlates with dimerization. *Nucleic Acids Res* **45**, 3378-3394, doi:10.1093/nar/gkx066 (2017).
210 An, P. *et al.* APOBEC3G genetic variants and their influence on the progression to AIDS. *J Virol* **78**, 11070-11076, doi:10.1128/JVI.78.20.11070-11076.2004

78/20/11070 [pii] (2004).

211 Reddy, K. *et al.* Functional characterization of Vif proteins from HIV-1 infected patients with different APOBEC3G haplotypes. *Aids* **30**, 1723-1729, doi:10.1097/qad.0000000000001113 (2016).
212 Duggal, N. K., Fu, W., Akey, J. M. & Emerman, M. Identification and antiviral activity of common polymorphisms in the APOBEC3 locus in human populations. *Virology* **443**, 329-337, doi:10.1016/j.virol.2013.05.016 (2013).
213 Belanger, K. & Langlois, M. A. Comparative analysis of the gene-inactivating potential of retroviral restriction factors APOBEC3F and APOBEC3G. *J Gen Virol* **96**, 2878-2887, doi:10.1099/vir.0.000214 (2015).
214 Pollack, R. A. *et al.* Defective HIV-1 Proviruses Are Expressed and Can Be Recognized by Cytotoxic T Lymphocytes, which Shape the Proviral Landscape. *Cell Host Microbe* **21**, 494-506.e494, doi:10.1016/j.chom.2017.03.008 (2017).
215 Kim, E. Y. *et al.* Human APOBEC3 induced mutation of human immunodeficiency virus type-1 contributes to adaptation and evolution in natural infection. *PLoS Pathog* **10**, e1004281, doi:10.1371/journal.ppat.1004281 (2014).
216 Kim, E. Y. *et al.* Human APOBEC3G-mediated editing can promote HIV-1 sequence diversification and accelerate adaptation to selective pressure. *J Virol* **84**, 10402-10405, doi:JVI.01223-10 [pii]

10.1128/JVI.01223-10 (2010).
217 Zennou, V. & Bieniasz, P. D. Comparative analysis of the antiretroviral activity of APOBEC3G and APOBEC3F from primates. *Virology* **349**, 31-40, doi:10.1016/j.virol.2005.12.035 (2006).
218 Chaipan, C., Smith, J. L., Hu, W. S. & Pathak, V. K. APOBEC3G restricts HIV-1 to a greater extent than APOBEC3F and APOBEC3DE in human primary CD4+ T cells and macrophages. *J Virol* **87**, 444-453, doi:JVI.00676-12 [pii]

10.1128/JVI.00676-12 (2013).
219 Sato, K. *et al.* APOBEC3D and APOBEC3F potently promote HIV-1 diversification and evolution in humanized mouse model. *PLoS Pathog* **10**, e1004453, doi:10.1371/journal.ppat.1004453 (2014).
220 Monajemi, M., Woodworth, C. F., Benkaroun, J., Grant, M. & Larijani, M. Emerging complexities of APOBEC3G action on immunity and viral fitness during HIV infection and treatment. *Retrovirology* **9**, 35, doi:1742-4690-9-35 [pii]

10.1186/1742-4690-9-35 (2012).
221 Monajemi, M. *et al.* Positioning of APOBEC3G/F mutational hotspots in the human immunodeficiency virus genome favors reduced recognition by CD8+ T cells. *PLoS One* **9**, e93428, doi:10.1371/journal.pone.0093428 (2014).
222 Squires, K. D., Monajemi, M., Woodworth, C. F., Grant, M. D. & Larijani, M. Impact of APOBEC Mutations on CD8+ T Cell Recognition of HIV Epitopes Varies Depending on the Restricting HLA. *J Acquir Immune Defic Syndr* **70**, 172-178, doi:10.1097/qai.0000000000000689 (2015).
223 Fourati, S. *et al.* E138K and M184I mutations in HIV-1 reverse transcriptase coemerge as a result of APOBEC3 editing in the absence of drug exposure. *AIDS* **26**, 1619-1624, doi:10.1097/QAD.0b013e3283560703 (2012).

224 Fourati, S. *et al.* HIV-1 genome is often defective in PBMCs and rectal tissues after long-term HAART as a result of APOBEC3 editing and correlates with the size of reservoirs. *J Antimicrob Chemother* **67**, 2323-2326, doi:dks219 [pii]

10.1093/jac/dks219 (2012).
225 Knoepfel, S. A. *et al.* Comparison of G-to-A mutation frequencies induced by APOBEC3 proteins in H9 cells and peripheral blood mononuclear cells in the context of impaired processivities of drug-resistant human immunodeficiency virus type 1 reverse transcriptase variants. *J Virol* **82**, 6536-6545, doi:JVI.00554-08 [pii]

10.1128/JVI.00554-08 (2008).
226 McCallum, M. *et al.* Basis for early and preferential selection of the E138K mutation in HIV-1 reverse transcriptase. *Antimicrob Agents Chemother* **57**, 4681-4688, doi:10.1128/aac.01029-13 (2013).
227 Noguera-Julian, M. *et al.* Contribution of APOBEC3G/F activity to the development of low-abundance drug-resistant human immunodeficiency virus type 1 variants. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases* **22**, 191-200, doi:10.1016/j.cmi.2015.10.004 (2016).
228 Russell, R. A., Moore, M. D., Hu, W. S. & Pathak, V. K. APOBEC3G induces a hypermutation gradient: purifying selection at multiple steps during HIV-1 replication results in levels of G-to-A mutations that are high in DNA, intermediate in cellular viral RNA, and low in virion RNA. *Retrovirology* **6**, 16, doi:1742-4690-6-16 [pii]

10.1186/1742-4690-6-16 (2009).
229 Ji, J. P. & Loeb, L. A. Fidelity of HIV-1 reverse transcriptase copying RNA in vitro. *Biochemistry* **31**, 954-958 (1992).
230 Refsland, E. W. *et al.* Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Res* **38**, 4274-4284, doi:10.1093/nar/gkq174 (2010).
231 Koning, F. A. *et al.* Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. *J Virol* **83**, 9474-9485, doi:10.1128/JVI.01089-09 (2009).
232 Dunn, L. L., Boyer, P. L., McWilliams, M. J., Smith, S. J. & Hughes, S. H. Mutations in human immunodeficiency virus type 1 reverse transcriptase that make it sensitive to degradation by the viral protease in virions are selected against in patients. *Virology* **484**, 127-135, doi:10.1016/j.virol.2015.05.020 (2015).
233 Mulder, L. C. *et al.* Moderate influence of human APOBEC3F on HIV-1 replication in primary lymphocytes. *J Virol* **84**, 9613-9617, doi:JVI.02630-09 [pii]

10.1128/JVI.02630-09 (2010).
234 Miyagi, E. *et al.* Stably expressed APOBEC3F has negligible antiviral activity. *J Virol* **84**, 11067-11075, doi:JVI.01249-10 [pii]

10.1128/JVI.01249-10 (2010).
235 Browne, E. P., Allers, C. & Landau, N. R. Restriction of HIV-1 by APOBEC3G is cytidine deaminase-dependent. *Virology* **387**, 313-321, doi:S0042-6822(09)00133-0 [pii]

10.1016/j.virol.2009.02.026 (2009).

236 Refsland, E. W., Hultquist, J. F. & Harris, R. S. Endogenous origins of HIV-1 G-to-A hypermutation and restriction in the nonpermissive T cell line CEM2n. *PLoS Pathog* **8**, e1002800, doi:10.1371/journal.ppat.1002800

PPATHOGENS-D-12-00221 [pii] (2012).
237 Desimmie, B. A. *et al.* APOBEC3 proteins can copackage and comutate HIV-1 genomes. *Nucleic Acids Res* **44**, 7848-7865, doi:10.1093/nar/gkw653 (2016).
238 Krisko, J. F., Martinez-Torres, F., Foster, J. L. & Garcia, J. V. HIV Restriction by APOBEC3 in Humanized Mice. *PLoS Pathog* **9**, e1003242, doi:10.1371/journal.ppat.1003242

PPATHOGENS-D-12-02677 [pii] (2013).
239 Sato, K. *et al.* Remarkable lethal G-to-A mutations in vif-proficient HIV-1 provirus by individual APOBEC3 proteins in humanized mice. *J Virol* **84**, 9546-9556, doi:JVI.00823-10 [pii]

10.1128/JVI.00823-10 (2010).
240 Sadler, H. A., Stenglein, M. D., Harris, R. S. & Mansky, L. M. APOBEC3G contributes to HIV-1 variation through sublethal mutagenesis. *J Virol* **84**, 7396-7404, doi:JVI.00056-10 [pii]

10.1128/JVI.00056-10 (2010).
241 Gao, L. *et al.* Apparent defects in processive DNA synthesis, strand transfer, and primer elongation of Met-184 mutants of HIV-1 reverse transcriptase derive solely from a dNTP utilization defect. *J Biol Chem* **283**, 9196-9205, doi:10.1074/jbc.M710148200 (2008).
242 Xu, H. T. *et al.* Compensation by the E138K mutation in HIV-1 reverse transcriptase for deficits in viral replication capacity and enzyme processivity associated with the M184I/V mutations. *J Virol* **85**, 11300-11308, doi:10.1128/jvi.05584-11 (2011).
243 Sharma, P. L. & Crumpacker, C. S. Decreased processivity of human immunodeficiency virus type 1 reverse transcriptase (RT) containing didanosine-selected mutation Leu74Val: a comparative analysis of RT variants Leu74Val and lamivudine-selected Met184Val. *J Virol* **73**, 8448-8456 (1999).
244 Gao, H. Q., Boyer, P. L., Arnold, E. & Hughes, S. H. Effects of mutations in the polymerase domain on the polymerase, RNase H and strand transfer activities of human immunodeficiency virus type 1 reverse transcriptase. *J Mol Biol* **277**, 559-572, doi:10.1006/jmbi.1998.1624 (1998).
245 Back, N. K. *et al.* Reduced replication of 3TC-resistant HIV-1 variants in primary cells due to a processivity defect of the reverse transcriptase enzyme. *EMBO J* **15**, 4040-4049 (1996).
246 Boyer, P. L. & Hughes, S. H. Analysis of mutations at position 184 in reverse transcriptase of human immunodeficiency virus type 1. *Antimicrob Agents Chemother* **39**, 1624-1628 (1995).
247 Mangeat, B. *et al.* Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature* **424**, 99-103 (2003).
248 Zhang, H. *et al.* The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* **424**, 94-98 (2003).

249 Harris, R. S., Petersen-Mahrt, S. K. & Neuberger, M. S. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Molecular cell* **10**, 1247-1253 (2002).
250 Wiegand, H. L., Doehle, B. P., Bogerd, H. P. & Cullen, B. R. A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *EMBO J* **23**, 2451-2458, doi:10.1038/sj.emboj.7600246

7600246 [pii] (2004).
251 Zheng, Y. H. *et al.* Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *J Virol* **78**, 6073-6076, doi:10.1128/JVI.78.11.6073-6076.2004

78/11/6073 [pii] (2004).
252 Kao, S. *et al.* The human immunodeficiency virus type 1 Vif protein reduces intracellular expression and inhibits packaging of APOBEC3G (CEM15), a cellular inhibitor of virus infectivity. *J Virol* **77**, 11398-11407 (2003).
253 Marin, M., Rose, K. M., Kozak, S. L. & Kabat, D. HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nature medicine* **9**, 1398-1403, doi:10.1038/nm946 (2003).
254 Mangeat, B., Turelli, P., Liao, S. & Trono, D. A single amino acid determinant governs the species-specific sensitivity of APOBEC3G to Vif action. *J Biol Chem* **279**, 14481-14483, doi:10.1074/jbc.C400060200

C400060200 [pii] (2004).
255 Schrofelbauer, B., Chen, D. & Landau, N. R. A single amino acid of APOBEC3G controls its species-specific interaction with virion infectivity factor (Vif). *Proc Natl Acad Sci U S A* **101**, 3927-3932, doi:10.1073/pnas.0307132101

0307132101 [pii] (2004).
256 Xu, H. *et al.* A single amino acid substitution in human APOBEC3G antiretroviral enzyme confers resistance to HIV-1 virion infectivity factor-induced depletion. *Proc Natl Acad Sci U S A* **101**, 5652-5657, doi:10.1073/pnas.0400830101

0400830101 [pii] (2004).
257 Zhang, W., Du, J., Evans, S. L., Yu, Y. & Yu, X. F. T-cell differentiation factor CBF-beta regulates HIV-1 Vif-mediated evasion of host restriction. *Nature* **481**, 376-379, doi:10.1038/nature10718 (2011).
258 Vartanian, J. P., Meyerhans, A., Sala, M. & Wain-Hobson, S. G-->A hypermutation of the human immunodeficiency virus type 1 genome: evidence for dCTP pool imbalance during reverse transcription. *Proc Natl Acad Sci U S A* **91**, 3092-3096 (1994).
259 Holmes, M., Zhang, F. & Bieniasz, P. D. Single-Cell and Single-Cycle Analysis of HIV-1 Replication. *PLoS Pathog* **11**, e1004961, doi:10.1371/journal.ppat.1004961 (2015).
260 Langlois, M. A., Beale, R. C., Conticello, S. G. & Neuberger, M. S. Mutational comparison of the single-domained APOBEC3C and double-domained APOBEC3F/G anti-retroviral cytidine deaminases provides insight into their DNA target site specificities. *Nucleic Acids Res* **33**, 1913-1923, doi:33/6/1913 [pii]

10.1093/nar/gki343 (2005).

261 Pollack, R. A. *et al.* Defective HIV-1 Proviruses Are Expressed and Can Be Recognized by Cytotoxic T Lymphocytes, which Shape the Proviral Landscape. *Cell Host Microbe* **21**, 494-506 e494, doi:10.1016/j.chom.2017.03.008 (2017).
262 Baig, T. T., Feng, Y. & Chelico, L. Determinants of efficient degradation of APOBEC3 restriction factors by HIV-1 Vif. *J Virol* **88**, 14380-14395, doi:10.1128/JVI.02484-14 (2014).
263 Ooms, M., Majdak, S., Seibert, C. W., Harari, A. & Simon, V. The localization of APOBEC3H variants in HIV-1 virions determines their antiviral activity. *J Virol* **84**, 7961-7969, doi:10.1128/JVI.00754-10 (2010).
264 Refsland, E. W. *et al.* Natural polymorphisms in human APOBEC3H and HIV-1 Vif combine in primary T lymphocytes to affect viral G-to-A mutation levels and infectivity. *PLoS Genet* **10**, e1004761, doi:10.1371/journal.pgen.1004761 (2014).
265 Albin, J. S. *et al.* A single amino acid in human APOBEC3F alters susceptibility to HIV-1 Vif. *J Biol Chem* **285**, 40785-40792, doi:M110.173161 [pii]

10.1074/jbc.M110.173161 (2010).
266 Smith, J. L. & Pathak, V. K. Identification of specific determinants of human APOBEC3F, APOBEC3C, and APOBEC3DE and African green monkey APOBEC3F that interact with HIV-1 Vif. *J Virol* **84**, 12599-12608, doi:JVI.01437-10 [pii]

10.1128/JVI.01437-10 (2010).
267 Zerbino, D. R. *et al.* Ensembl 2018. *Nucleic Acids Res* **46**, D754-D761, doi:10.1093/nar/gkx1098 (2018).
268 Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539, doi:10.1038/msb.2011.75 (2011).
269 Rose, P. P. & Korber, B. T. Detecting hypermutations in viral sequences with an emphasis on G --> A hypermutation. *Bioinformatics* **16**, 400-401 (2000).
270 Dang, Y. *et al.* Identification of a single amino acid required for APOBEC3 antiretroviral cytidine deaminase activity. *J Virol* **85**, 5691-5695, doi:10.1128/jvi.00243-11 (2011).
271 Albin, J. S., Hache, G., Hultquist, J. F., Brown, W. L. & Harris, R. S. Long-term restriction by APOBEC3F selects human immunodeficiency virus type 1 variants with restored Vif function. *J Virol* **84**, 10209-10219, doi:JVI.00632-10 [pii]

10.1128/JVI.00632-10 (2010).
272 Siu, K. K., Sultana, A., Azimi, F. C. & Lee, J. E. Structural determinants of HIV-1 Vif susceptibility and DNA binding in APOBEC3F. *Nat Commun* **4**, 2593, doi:10.1038/ncomms3593 (2013).
273 Bohn, M. F. *et al.* Crystal structure of the DNA cytosine deaminase APOBEC3F: the catalytically active and HIV-1 Vif-binding domain. *Structure (London, England : 1993)* **21**, 1042-1050, doi:S0969-2126(13)00122-6 [pii]

10.1016/j.str.2013.04.010 (2013).
274 Jiang, X., Buxbaum, J. N. & Kelly, J. W. The V122I cardiomyopathy variant of transthyretin increases the velocity of rate-limiting tetramer dissociation, resulting in accelerated amyloidosis. *Proceedings of the National Academy of Sciences* **98**, 14943-14948, doi:10.1073/pnas.261419998 (2001).
275 Nelson, D., Lehninger, Albert L, & Cox, Michael M. *Lehninger principles of biochemistry*. 5th ed edn, (W.H. Freeman., 2008).

276 Meyerson, N. R. & Sawyer, S. L. Two-stepping through time: mammals and viruses. *Trends Microbiol* **19**, 286-294, doi:10.1016/j.tim.2011.03.006 (2011).
277 Naldini, L. *et al.* In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263-267 (1996).
278 Kao, S. *et al.* Production of infectious human immunodeficiency virus type 1 does not require depletion of APOBEC3G from virus-producing cells. *Retrovirology* **1**, 27, doi:10.1186/1742-4690-1-27

1742-4690-1-27 [pii] (2004).
279 Holmes, R. K., Koning, F. A., Bishop, K. N. & Malim, M. H. APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. *J Biol Chem* **282**, 2587-2595, doi:M607298200 [pii]

10.1074/jbc.M607298200 (2007).
280 Artimo, P. *et al.* ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res* **40**, W597-603, doi:10.1093/nar/gks400 (2012).